이학박사 학위논문

Functional brain connectivity of audiovisual speech processing using graph filtration

시청각 언어 처리시의 기능적 뇌 연결성: 그래프 필트레이션 방법 적용

2012년 8월

서울대학교 대학원

협동과정 인지과학 전공

김 희 정

Functional brain connectivity of audiovisual speech processing using graph filtration

시청각 언어 처리시의 기능적 뇌 연결성:

그래프 필트레이션 방법 적용

지도 교수 이 동 수

이 논문을 이학박사 학위논문으로 제출함

2012년 7월

서울대학교 대학원

협동과정 인지과학 전공

김 희 정

김희정의 이학박사 학위논문을 인준함

2012년 7월

위 원 장	장 병 탁	(인)
부위원장	이 동 수	(인)
위 원	김 청 택	(인)
위 원	정 민 화	(인)
위 원	강 은 주	(인)

Functional brain connectivity of audiovisual speech processing using graph filtration

Kim, Heejung

Interdisciplinary Program in Cognitive Science

The Graduate School

A Thesis Submitted to the Faculty of Interdisciplinary Program in Cognitive Science, in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Science at the Seoul National University, Seoul, Korea

July 2012

Approved by thesis committee:

Chairman	**Byoung-Tak Zhang**
Vice-chairman	**Dong Soo Lee**
Member	**Cheongtag Kim**
Member	**Minhwa Chung**
Member	**Eunjoo Kang**

Abstract

Functional brain connectivity of audiovisual speech processing using graph filtration

Heejung Kim

Interdisciplinary Program in Cognitive Science

The Graduate School

Seoul National University

Several brain regions have been implicated in audiovisual speech integration. However, the functional network that exists between these regions to facilitate multisensory speech integration remains unclear. Previous studies suggest that the superior temporal sulcus/gyrus (STS/STG) is a critical brain area for multisensory integration, but functional connectivity between STS/STG and other regions involved in audiovisual speech processing has not been observed consistently, perhaps in part due to differences in the kinds of tasks or thresholds used across studies. To avoid this issue, the current study used graph filtration in a persistent homology framework to investigate the functional networks among brain areas activated during audiovisual speech

processing. The audiovisual speech task in this study included 4 conditions that differed in the sensory modality used to deliver speech: audiovisual speech (AV), auditory speech (A), visual speech (V), and audiovisual non-speech (C). I constructed three speech-specific brain networks that were represented by barcode and minimum spanning tree using single linkage distance and compared functional networks. These results revealed that the audiovisual speech network was focused on connectivity between the posterior STG (pSTG) and inferior frontal region, while the visual speech network was focused on the connectivity between the hippocampal and inferior frontal regions. Prominent functional connectivity was observed between the pSTG and frontal regions during audiovisual speech integration, and functional connectivity between the temporal regions and auditory or visual areas was driven by auditory or visual speech modalities. Additionally, meaningful visual speech without auditory speech seems to depend on the retrieval of previous memories related to speech production. The results of this study suggest that the mechanisms by which the brain uses visual information to improve auditory speech perception may involve fronto-temporal connectivity with motor areas related to speech production, rather than visual sensory areas.

Keywords: Functional brain connectivity; Audiovisual speech integration; Task fMRI; Persistent homology; Graph filtration

Student Number: 2005-30932

To my family,

CONTENTS

List of Tables

List of Figures

represented by barcodes. (e) Barcodes were rearranged as a single linkage dendrogram according to the node index. (f) Algebraic matrix was formed as a single linkage matrix.

Figure 5. Areas with significantly increased activation, displayed through whole brain analysis for each condition compared to control. Brain regions in which BOLD signal changes were observed are shown in red-yellow color. (uncorrected $p < 0.005$, extend threshold $k > 10$).

Figure 6. The connectivity matrix. (a) Each connectivity map represents the positive correlation matrix in each condition. (b) Connectivity maps were represented at uncorrected $p < 0.01$. (c) Each connectivity map represents the single linkage matix. The heatmap indicates shorter distance between nodes, with yellow indicating higher correlation. (d) Brain regions were selected from GLM analysis for functional connectivity. All brain regions are described in Table 3.

Figure 7. Differences in the functional connectivity between different speech conditions. (a) Connectivity difference maps were computed between correlation matrices using nonparametric permutation testing at $p < 0.05$. (b) Differences between single linkage matrices (SLMs) were also computed using nonparametric permutation testing at $p < 0.05$.

Figure 8. Barcodes representing functionally connected components during the three speech conditions across a range of filtration values, ε. (a) The vertical axis displays the number of connected component and the horizontal axis shows filtration value. Red barcodes represent the

AV condition, green for A, and blue for V conditions. (b) Differences in the slope of barcodes between conditions display. For differences of functional connectivity through filtration values, nonparametric permutation testing was performed and significance was defined as p values ($p < 0.05$). AV, A and V represent audiovisual, auditory, and visual conditions, respectively.

Figure 9. The connected component matrix through filtration values. (a) Barcodes of connected component and single linkage dendrograms of speech conditions (b) Connected component matrix through filtration values, in which filtration values were 0.15, 0.20, 0.25, 0.35, 0.40, 0.50 and 0.70.

Figure 10. The minimum spanning trees (MSTs) of three speech conditions. (a) The MSTs formed a circular pattern. Color displays brain regions, such as black, yellow, brown, and light blue for frontal, temporal, subcortical, and occipital regions, respectively. (b) The MSTs were illustrated in the corresponding brain structure. Node colors were as follows: red, green, blue, and magenta for frontal, temporal, subcortical, and occipital regions, respectively. Common edge was displayed as a red line on the circular pattern.

Figure 11. Distance and single linkage matrices. In the top panel (a), the distance matrices were obtained by 1 - correlation after estimation between 90 regions of interests. In the middle panel (b), each correlation matrix was thresholded at an arbitrary value (correlation coefficient > 0.7). Correlations in two diagonal lines represent correlations between homologous regions in the bilateral hemisphere.

In the bottom panel (c), the single linkage matrices were obtained based on the distance matrix. The heatmap denotes sequence of component merging during analysis, with brighter colors representing earlier merging. Yellow color represents higher correlation. Correlations along the two diagonal lines represent correlations both between regions within the same hemisphere and between homologue regions in the bilateral hemisphere.

Figure 12. Connected components of brain network revealed using graph filtration. (a) Barcodes representing connected components through varying filtration values during each condition. (b) Single linkage dendrograms of each condition represented by barcodes rearranged according to node index.

Supplementary Figure S1. Differences in correlation matrices between conditions. (a) Connectivity differences were found by using nonparametric permutation testing between audiovisual vs. control and auditory vs. control at uncorrected $p < 0.001$ (b) The significance in three comparisons: audiovisual vs. auditory, audiovisual vs. visual, and auditory vs. visual at uncorrected $p < 0.01$. White rectangle denotes statistically significant connection.

Supplementary Figure S2. The connected component matrix through filtration values. Each row represents filtration values, 0.10, 0.16, 0.20, 0.22, 0.28 and 0.30 respectively. Each column represents the audiovisual (AV), auditory (A), visual (V), and control (C) conditions, respectively.

Supplementary Figure S3. The minimum spanning trees (MSTs) of all conditions. (a) MSTs represented in a circular pattern. Colors denote brain region, with black, yellow, brown, light blue, blue, and pink for frontal, subcortical, parietal, temporal, basal ganglia, and occipital regions, respectively. (b) MSTs illustrated in the corresponding brain structure. Node colors are as follows: red, blue, yellow, magenta, green, and light blue for frontal, subcortical, parietal, temporal, basal ganglia, and occipital regions, respectively. (c) Common MSTs. Common edges consisted of edges between homologous regions in bilateral hemispheres and between neighboring regions, with a few exceptions.

Supplementary Figure S4. The minimum spanning trees (MSTs) of all conditions. (a) Each MST subtracted from the common MST. (b) Common MSTs represented in a circular pattern.

1. Introduction

1. 1 Audiovisual speech integration

Understanding speech in noisy environments is a difficult task or needs many efforts. How humans understand speech in a noisy environment has become an important issue. In these instances, additive information from the speaker's visual speech can facilitate auditory speech processing. Combining information from the auditory and visual modalities (congruent lip movement and auditory speech) reduces noise and enhances speech perception. Multisensory integration may partly explain this phenomenon. However, when the visual information has an incongruent or ambiguous relationship with auditory speech, a perceptual illusion known as the McGurk effect may arise in which the incongruous audiovisual information forms an integrated percept of the two modalities (McGurk and MacDonald, 1976).

In face-to-face conversation, speech perception involves processing of auditory speech sounds as well as visual speech, such as mouth movements. Listeners are typically not conscious of the visual information during conversation and instead focus on auditory speech information. When visual lip or mouth information is presented in the

absence of an auditory speech stimulus, normal subjects have great difficulty understanding the speech (although visual speech information is important for hearing-impaired individuals who can learn to lip-read to take advantage of visual cues). Sekiyama et al., (2003) reported that the left superior temporal sulcus (STS) was involved in cross-modal binding of audiovisual speech perception using the McGurk effect paradigm. Previous studies have demonstrated a critical role of the left STS in the perception of multisensory speech as well as non-speech stimuli (Calvert et al., 2000; Wright et al., 2003; Beauchamp et al., 2005; Miller and D'Esposito, 2005; Kang et al., 2006; Nath and Beauchamp, 2011). Many neuroimaging studies have reported that the superior temporal gyrus (STG) and STS are also activated during lip-reading (Olson et al., 2002; Calvert et al., 1997).

Anatomical studies in nonhuman primates have shown that the STS is connected with both auditory and visual cortices (Lewis and Van Essen, 2000). Semantic or syntactic violations in speech comprehension activate fronto-temporal circuits (Friederici et al., 2000, 2003). Many regions are involved in audiovisual speech perception, including the STS, Broca's area, visual cortex, auditory cortex, and supramarginal gyrus. Perceptual binding during multisensory integration emerges through cooperation of numerous distinct brain

regions. In audiovisual speech perception, the STS/STG is thought to be involved in multisensory integration. It is unknown whether perceptual integration is processed in a hierarchical manner, in which simple to complex information processing is relayed from sensory cortices to higher-level sensory and associative cortices. Understanding the connectivity between the structures involved in multisensory integration may help to explain how processing of speech information is handled in the nervous system. The reason I want to investigate connectivity may be here. Investigating these connections can be difficult because co-activation of different brain regions does not necessarily imply functional connectivity between the regions. Little is known about the functional connectivity for multisensory integration in the network of task-relevant brain structures during audiovisual speech perception.

1. 2 Functional brain networks during task fMRI

Interactions between large-scale neural systems can be indirectly estimated using blood oxygenation level-dependent (BOLD) functional magnetic resonance imaging (fMRI). Traditionally, fMRI studies have focused on brain regions showing task-related increases in neural activity or attentional increases in neural activity during resting

state (Raichle et al., 2001). In past decades, research has been focused on aspects of brain function based on functional integration or segregation (Frackowiak et al., 2004). Such studies have identified brain regions that are correlated with regressors of interest or have investigated interactivity of specific regions with the rest of the brain. Recently, fMRI has been used as a powerful tool for studying large-scale functional connectivity. This approach identifies interregional interactions using correlation between the fMRI time-series of distinct brain regions. Functional connectivity, which is constructed through correlations between signals of brain regions, is thought to represent the functional architecture of the brain, yielding fundamental insights into brain organization (Bullmore and Sporns, 2009; Smith et al., 2009). The identification of functional networks using fMRI relies on the assumption that correlations in BOLD signals are indicative of connectivity (Figure 1). However, since fMRI is an indirect measure of neural activity, these signals can be confounded by physiological noise such as cardiac, respiratory, and head motion.

The identification of functional brain networks is a major goal of neuroimaging studies. Recently, many neuroimaging studies have focused on the relationships between distinct brain areas, described as brain connectivity. Brain connectivity describes networks of brain

regions connected by functional associations or by anatomical connections. Identifying anatomical networks in the brain can help us to understand the fundamental architecture of interregional connections, while characterizing functional networks will help us to understand how this architecture supports neurophysiological dynamics.

Brain networks consist of nodes and edges, where nodes are anatomically or functionally predefined brain regions, and edges represent the relationship (such as correlation) between these regions. This framework enables graphical representation of brain functional networks. The organization of brain networks can be observed at three different levels. First, the relationship between single neurons is classified as the microscale organization. Second, macroscale organization refers to the relationship between anatomically distinct brain regions. Last, the relationship between neuronal groups or populations is known as the mesoscale organization (Sporns et al., 2005).

Correlation-based methods including the Pearson correlation are widely used in capturing network connections (Smith et al., 2011). In general, correlation-based measures of connectivity between distinct brain areas do not specify directionality. Directionality in the connection between two regions is inferred to imply causality between

the activities in the two regions. This distinction between correlations and direct functional connections is sometimes referred to as functional versus effective connectivity (Friston, 1994). Directionality is more difficult to estimate than the existence of connections. Effective connectivity based on directionality estimates the causal influence that each element of a system exerts on other elements. Functional connectivity is defined as the undirected association or correlation between neural activities recorded from anatomically separated regions (Friston et al. 1994). The concept of functional connectivity refers to the statistical association between spatially distributed brain activation time series (Lee et al., 2006).

Methods for functional connectivity analysis include voxel-wise correlations, partial least square analysis, and independent component analysis (Morris et al., 1999; McIntosh and Lobaugh, 2004; Grady et al., 2003). Studies of pair-wise connectivity throughout the brain have revealed functional networks related to various cognitive processes or individual variation (Smith et al., 2011). The characterization of brain connectivity has enabled us to understand the associations between the brain and human behavior.

To date, most brain network modeling using graph theory has been based on resting-state fMRI data. Many studies have shown

ongoing interconnections or highly correlated activity between multiple brain regions during the resting state (Smith et al., 2009; Buckner et al., 2009; Greicius, 2009; Archard et al., 2006; Salvador et al., 2005). Even in the absence of a specific task, spontaneous fluctuations in the fMRI BOLD signal were consistently detected (Damoiseaux et al., 2006). In addition, fMRI analysis of connectivity has clinical applications for detecting changes in connectivity between normal subjects and patients with neurological or psychological disorders (Caeyenberghs et al., 2012; Fornito et al., 2011; Lord et al., 2011). Topological metrics have been posited to be specific biomarkers for neurodegenerative diseases such as Alzheimer's disease or Schizophrenia (Supekar et al., 2008; Fornito et al., 2011). Several studies have described the changes in functional network topologies of cognitive tasks. Such changes have been evaluated using electroencephalography (EEG) recording during changes in working memory load with an n-back task (Pachou et al., 2008). Other studies have investigated this phenomenon using magnetoencephalography (MEG) (Bassett et al., 2009) or fMRI (Salvador et al., 2008; Ginestet et al., 2011a). Pantazatos et al. (2012) investigated condition-dependent functional networks in order to discriminate between the different subconscious processing of cognitive-emotional states within individual subjects.

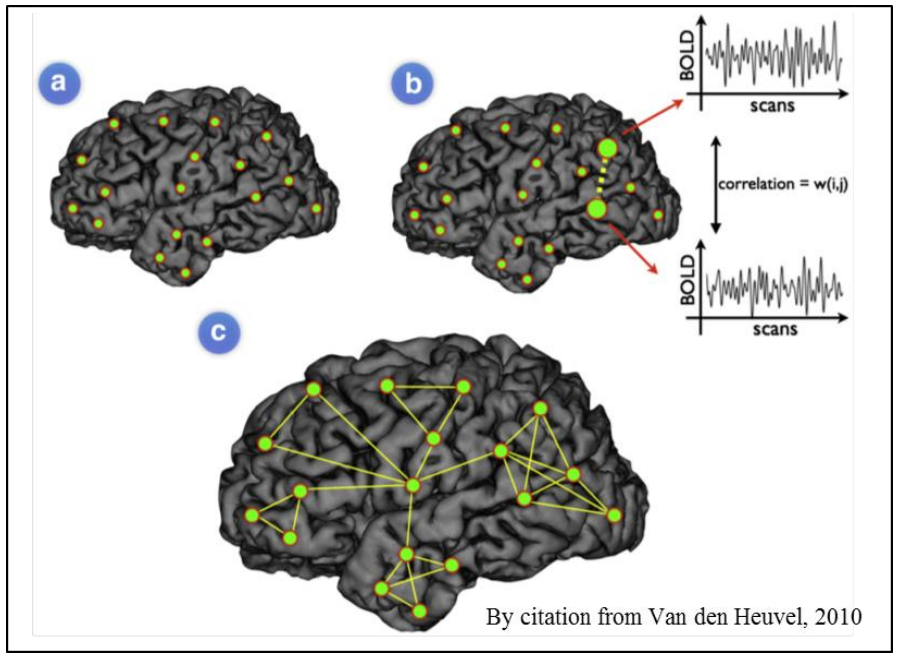


Figure 1. Schematic representation of the fMRI time-series used to identify functional networks. The connected network is represented as a graph, with nodes and edges. (a) Nodes denote brain regions, identified by a predefined template or activated regions. (b) Links representing the level of co-activation between brain regions, measured as correlation. (c) If co-activation is observed between two nodes above a given threshold, a functional connection is identified between the regions (Van den Heuvel, 2010).

1.3 Brain network analysis in persistent homology

Brain connectivity analysis has been developed through complex network analysis, a network analysis method with mathematical origins. Graph theory has proven to be a very informative method for exploring the relationship between brain function and human behavior (Bullmore and Sporns, 2009; Nakamura et al., 2009; Stam, 2010). Most studies using graph theory use a comparison with topological measures to investigate functional brain networks. Measures of functional network connectivity can be represented in multiple ways, and each measure reflects different topological aspects of connectivity. Some measures are represented as key graph properties, including the clustering coefficient, characteristic path length, node degree, node distribution, centrality, and modularity (Reijneveld et al., 2007; Sporns et al., 2004; Stam and Reijneveld, 2007). Of these metrics, a high clustering coefficient and the shortest path length characterize small-world network. Small-world network combines high levels of local and global efficiency. Graph theory thus, offers a broad selection of measures to examine and quantify the relationships between distinct brain regions.

However, most studies have focused on only a few measures, such as clustering coefficient, network path length, small-worldness, modularity, or connectivity degree. Although these measures based on graph theory give information about topological properties, the optimal measure for identifying and describing brain networks is not known. This poses a problem for the analysis and interpretation of neural functional network data (Figure 2). In network modeling, selecting an optimal threshold is an important issue. There are many ways to select thresholds, such as by statistical significance or by fixing the graph metrics. However, these methods produce somewhat arbitrary results. Figure 2 shows differences in the brain network identified by varying threshold (Lee et al., 2011b).

In this study, I applied new analytic methods to fMRI data without fixing a threshold. Using a persistent homology framework allowed me to observe the overall changes in the topological features across a range of thresholds. Bassett et al. (2012) showed that weak connections are strongly correlated with attention, memory, and negative symptoms scores. As such, these weak (and potentially difficult to identify) connections could be used as a potential biomarker. With an arbitrary threshold, weak connections are not represented in the network. For this reason, I applied graph filtration methods in the

persistent homology framework to investigate overall changes in the connectivity pattern instead of choosing particular measures or selecting an arbitrary threshold (Adler et al., 2010; Edelsbrunner et al., 2008; Lee et al., 2011a, 2011b and 2011c).

Graph filtration is a graph simplification technique that iteratively builds nested subgroups from the original graph (Figure 3). In the persistent homology framework, graph filtration was applied using features such as finite connected components to define the shape of the network. That is, persistent homology can observe the changes of topology during filtration and represent the topological features in an algebraic form. Using this formalism, topological features represented as connected components, are referred to as the 0^{th} Betti number. Herein, I define the shape of brain networks as a sequence of nested subgraphs using graph filtration. To apply the graph filtration, I computed distance measures, (i.e., 1- correlation values) in order to iteratively build nested subgroups of the original graph by piecing and forming patches of locally connected nodes. The connected graph is then transformed into an algebraic form called the single linkage distance, and the single linkage matrix (SLM) and minimum spanning tree (MST) are extracted from the single linkage distance.

Single linkage distance or single linkage clustering is a method of calculating distances between clusters, referred to as patches of locally connected nodes. The distance between two clusters is computed as the distance between the two closest elements in the two clusters. In single linkage distance, even though many nodes in each cluster may be very distant to each other, clusters may be forced together due to single nodes being close to each other. This concept is equivalent to MST. Single linkage distance can be represented as a single linkage dendrogram (SLD). This method successfully avoids the pitfalls associated with the selection of arbitrary thresholds and topological measures.

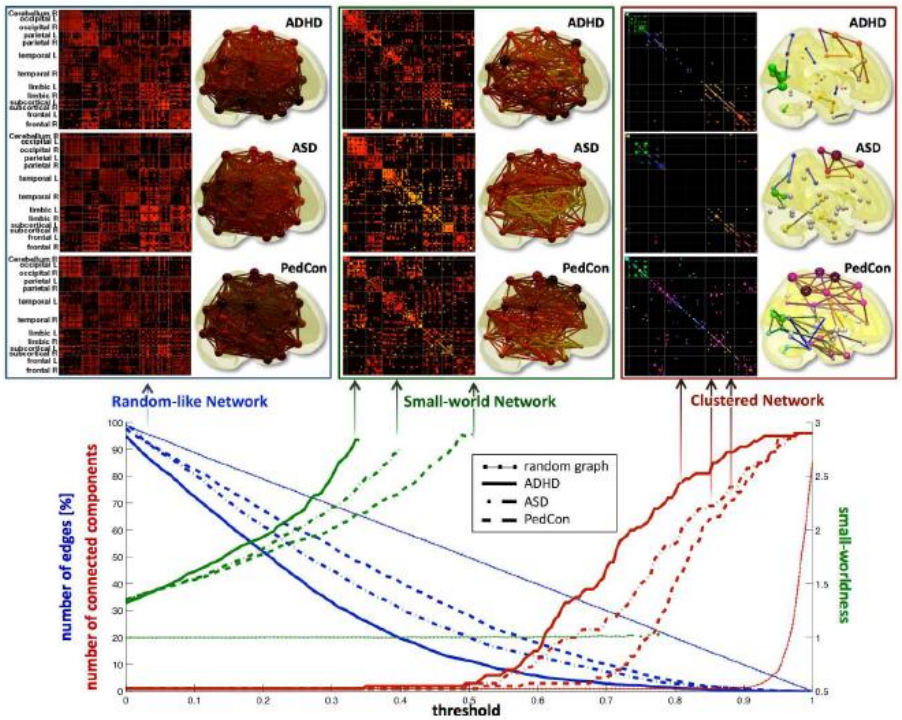


Figure 2. Different topological structures identified by using different thresholds. Threshold selection can bias the identified brain network toward being random-like, small-world, or clustered. Illustrated by Lee et al., 2011b.

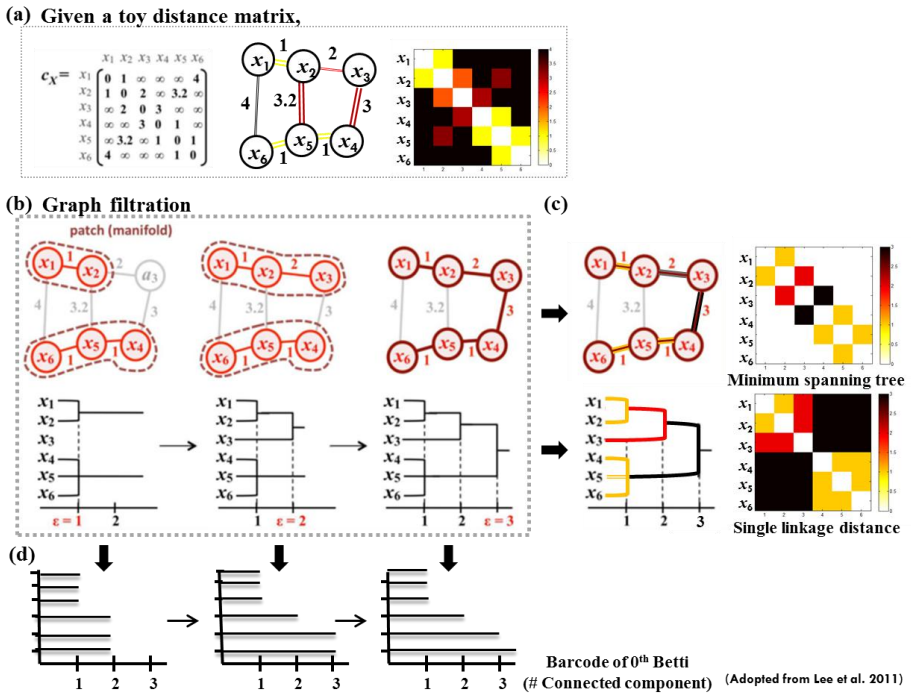


Figure 3. Schematic overview of graph filtration. (a) In a network of nodes, X = {x_1,..., x_6} and the distance *Cx* = 1 - correlation. (b) Graph filtration algorithm for the graph. (c) The resulting shape of the network can be equivalently represented using the single linkage distance matrix and the minimum spanning tree. (d) Barcode representations of 0^{th} Betti number (i.e., the number of connected components).

1.4 Purpose of this study

Cross-modal speech integration has been studied using a wide variety of methods. Regions involved in cross-modal speech integration have been identified using standard generalized linear modeling. In this study, I investigated how network organization between functionally connected brain regions differs depending on the type of audiovisual speech stimulus being processed. Although distributed brain regions can be co-activated, the relationship between them may not arise from a functional connection. I used audiovisual speech stimuli to investigate task-dependent functional networks during speech perception. Audiovisual speech information was presented together or alone from each sensory modality in normal subjects.

A graph filtration approach was used to investigate differences in functional networks during audiovisual, auditory, and visual speech processing. Using this approach, defining nodes remains somewhat challenging. For whole brain functional network, a priori template-based parcellation of the brain has been used to define the network nodes. Therefore, the nodes typically include brain regions with little or no task-related activation (Nakamura et al., 2009). However, recent studies have reported an overall correspondence in functional architecture between resting and task-related brain states (Molinari et

al., 2012; Smith et al., 2009). In the present study, I constructed functional networks using predefined brain regions and task-specific brain regions independently. For task-specific networks, I used region of interests (ROIs) selected from task-related activation through analysis of statistical parametric mapping. I investigated differences in the functional networks constructed between speech-related brain regions during audiovisual speech perception. In particular, my analysis focuses on superior temporal regions that have been implicated in audiovisual speech integration. I intend to identify functional connectivity of the region related to other sensory areas in task-specific network.

2. Materials and methods

2.1 Participants

In this study, twelve subjects (5 men and 7 women; mean age = 24.91 years; SD = 5.45 years) without any neurological or sensory disorder participated in the fMRI experiments. These subjects had normal or corrected-to-normal vision, and all were right-handed according to a modified Edinburgh Inventory (Oldfield, 1971). They were native Korean speakers, and had no experience in lip-reading. They received monetary compensation for their participation and provided written informed consent before the experiment. The fMRI procedure used in this study complied with the ethical guidelines of the local ethics committee (Samsung Seoul Hospital).

2.2 Cognitive task

I used the same experimental design as the one used in a previous PET study (Kang et al., 2006). In detail, the block design paradigm was used to generate four conditions. In all conditions, audiovisual moving clip stimuli were the inputs. The stimuli were

always bimodal consisting of visual face and auditory sounds. The auditory speech stimuli consisted of simple sentences in a male voice, and the visual speech stimuli comprised of congruent mouth movements with auditory speech. The sentences were of 3 to 4 units length in Korean (e.g., “A fish lives in the water” or “A sister is male”). A face with unopened mouth movements like chewing a gum was the visual non-speech stimuli, and white noises comprised the auditory non-speech stimuli. The four conditions were as follows. In the audiovisual speech condition (AV), both speech sounds and congruent mouth movement were presented as bimodal speech cues. In the auditory speech condition (A), speech sounds were delivered along with the visual non-speech stimuli. In the visual speech condition (V), visual speech cues were presented with the auditory non-speech stimuli. In the non-speech control (C) condition (C), auditory and visual non-speech stimuli were presented. During the AV, A and V speech conditions, subjects were required to press a button if they thought that the sentence they saw or heard was plausible in semantics. The probability of occurrence of true sentences was 50%. In the C conditions, the subjects were required to press the button on trial alternatively in order to maintain attention. Subjects performed the task twice (two scans); each scan consisted of five blocks (including the

fixation condition) and was repeated four times for each condition. Each scan lasted for 9 min 45 s, which included 15 s for processing dummy images.

2.3 fMRI data acquisition and preprocessing

FMRI data was acquired on a GE 1.5 T MRI scanner using a standard whole head coil. I obtained 195 T2* weighted echo-planar imaging volumes in each of the 20 axial planes parallel to the anterior and posterior commissure line (5 mm thick without gap, field of view [FOV] = 20 cm, 64 x 64, repetition time [TR] = 3000ms, echo time [TE] = 60ms, and flip angle = 90°). A high- resolution T1-weighted spoiled grass gradient recalled (SPGR) 3D MRI sequence (24 cm FOV, 124 slices) was obtained for anatomical reference. The initial preparation for functional network analysis involved passing all fMRI time series data through several preprocessing steps using Statistical Parametric Mapping (SPM 8, University College of London, UK), implemented in MATLAB 7.4 (Mathworks Inc., USA). The first five volumes of T2 functional images were discarded for head motion correction. The remaining images were realigned to the first image in the times series using affine registration with 6 parameters for removal

of movement artifacts. After realignment, the functional images were spatially co-registered to the T1-weighted anatomical image. Next, the functional images were spatially normalized into stereotaxic space using a nonlinear algorithm [Montreal Neurological Institute coordinates (MNI)] and then spatially smoothed (8 mm full-width at half-maximum Gaussian kernel). Additional preprocessing steps for mean time series extraction of connectivity networks was performed as per the standard method (Fornito et al., 2011; He et al., 2008, 2009 and 2010; Salvador et al., 2005). After smoothing, fMRI time series images were regressed out using 6-rigid body motion parameters and were subjected to linear detrending to correct for signal fluctuations during non-task condition. The residuals were further filtered at a frequency range of 0.01 to 0.08 Hz using band-pass filtering. This frequency range has shown to yield robust networks for functional connectivity (Fornito et al., 2010; Achard et al., 2006; Cordes et al., 2000).

2.4 Construction of brain functional connectivity

Nodes and edges were identified for the functional connectivity. Two types of ROIs were used to identify nodes. For task-specific functional connectivity, activated regions from a generalized linear

model (GLM) analysis were used to identify ROIs. For whole brain functional connectivity, predefined ROIs were used as nodes. Each node was assigned a value equal to the mean signal intensity of voxels included in the ROI. Pearson correlation values between nodes were used to identify edges and generate the functional connectivity matrix. The correlation values between ROIs reflected intrinsic correlations as well as correlated hemodynamic responses of pairs of regions evoked by each condition. The overall procedure for identifying functional brain connectivity is shown in Figure 4.

Schematic overview of the functional connectivity analysis

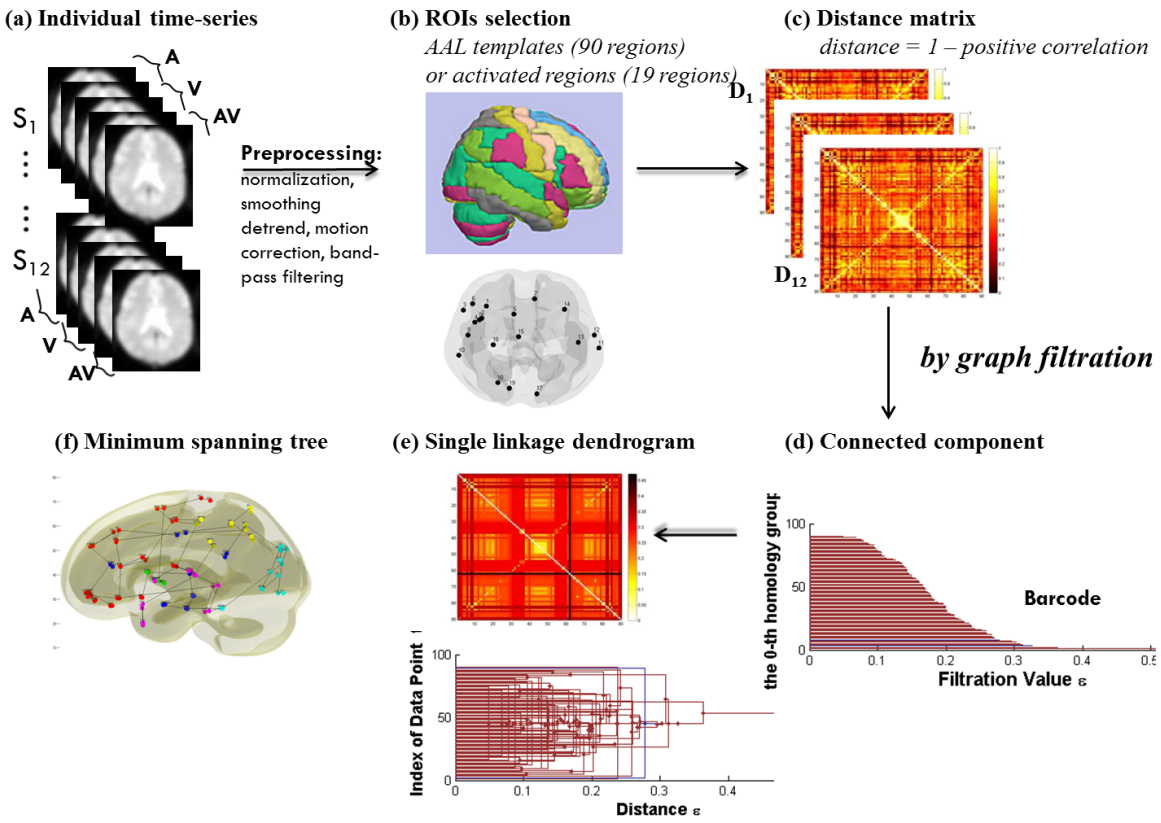


Figure 4. Schematic overview of the functional connectivity analysis. (a) The preprocessed, individual time-series data were divided into four conditions, and mean time-series were extracted from predefined regions. (b) Predefined regions became nodes in the network. (c) The correlation matrix and distance matrix between the defined regions were computed in each condition. (d) Connected components were represented by barcodes. (e) Barcodes were rearranged as a single linkage dendrogram according to the node index. Algebraic matrix was formed as a single linkage matrix. (f) Single linkage distance was also represented minimum spanning tree.

2.4.1 Defining nodes: Regions of interests

The first step in the functional connectivity analysis was to define nodes for the connectivity matrix. After defining ROIs for nodes in network, the value of a node was set as the mean BOLD signal intensity extracted from the defined ROI during the fMRI scan. From each fMRI image, we extracted mean BOLD signal values by averaging voxels within each of the defined ROIs. The first two images of each scan were discarded due to the lag in the hemodynamic response. The mean BOLD signal values were separated according to each of the four conditions.

Table 1. Whole-brain regions-of-interest (ROIs) in this analysis, defined by the Automated Anatomical Label (AAL) atlas and sorted by lobes. ROIs 1–45 are in the right hemisphere, while ROIs 45–90 are in the left hemisphere.

Regions	ROI NO. (R/L)	AAL Name	Abbreviations	BA
Frontal	1/90	Olfactory cortex	OC	
	2/89	Gyrus rectus	GR	BA11
	3/88	Inferior frontal gyrus, orbital part	F3O	BA47
	4/87	Inferior frontal gyrus, opercular part	F3OP	BA44
	5/86	Inferior frontal gyrus, triangular part	F3T	BA45
	6/85	Middle frontal gyrus, orbital part	F2O	BA10
	7/84	Middle frontal gyrus	F2	BA10
	8/83	Superior frontal gyrus, orbital part	F1O	BA6
	9/82	Superior frontal gyrus, medial orbital	F1MO	BA6
	10/81	Superior frontal gyrus, medial	F1M	BA6,8
	11/80	Superior frontal gyrus, dorsolateral	F1	BA6,8
	12/79	Paracentral lobule	PCL	BA4 (BA5, 6)
	13/78	Supplementary motor area	SMA	BA5
	14/77	Precentral gyrus	PRE	BA4
	15/76	Rolandic operculum	RO	BA43
Limbic	16/75	Anterior cingulate and paracingulate gyri	ACIN	BA24
	17/74	Median cingulate and paracingulate gyri	MCIN	BA23
	34/57	Thalamus	THA	
	35/56	Amygdala	AMYG	
	36/55	Hippocampus	HIP	
	37/54	Parahippocampal gyrus	PHIP	BA36
	38/53	Posterior cingulate gyrus	PCIN	BA29/30
Parietal	18/73	Postcentral gyrus	POST	BA1,2,3
	19/72	Superior parietal gyrus	P1	BA7
	20/71	Precuneus	PQ	BA7
	21/70	Inferior parietal, but supramarginal and angular gyri	P2	BA40

	22/69	Supramarginal gyrus	SMG	BA40
	23/68	Angular gyrus	AG	BA39
Temporal	24/67	Superior temporal gyrus	T1	BA22
	25/66	Heschl gyrus	HES	BA41
	26/65	Middle temporal gyrus	T2	BA21
	27/64	Inferior temporal gyrus	T3	BA20
	28/63	Temporal pole: superior temporal gyrus	T1P	BA38
	29/62	Temporal pole: middle temporal gyrus	T2P	BA38
	30/61	Insula	IN	BA13, 14
Basal ganglia	31/60	Caudate nucleus	CAU	BA 25
	32/59	Lenticular nucleus, putamen	PUT	
	33/58	Lenticular nucleus, pallidum	PAL	
Occipital	39/52	Fusiform gyrus	FUSI	BA37
	40/51	Inferior occipital gyrus	O3	BA18
	41/50	Middle occipital gyrus	O2	BA19
	42/49	Superior occipital gyrus	O1	BA19
	43/48	Calcarine fissure and surrounding cortex	V1	BA17
	44/47	Cuneus	Q	BA18
	45/46	Lingual gyrus	LING	BA 17, 18, 19

2.4.2 Edges: interregional correlation matrix

The second step of the functional connectivity analysis was to construct correlation maps across conditions. To construct correlation maps, Pearson correlations between the mean time-series BOLD signals extracted from the ROIs were computed for each condition in each subject. This resulted in {90 × 90} of whole brain connectivity for each condition and each subject and {19 × 19} matrices of task-related connectivity (Figure 4 (c)). In this study, I used only positive correlation values, and negative correlation values were set to zero due to the unknown biological meaning of negative correlation between brain regions.

The connectivity maps of each condition comprised of 12 correlation matrices. These weighted positive correlation matrices were used to analyze the functional networks. I performed Fisher's r-to-z transformation on each correlation matrix by $z = 1/2 \log [(1 + r) / (1 - r)]$, and then the z matrices were averaged across the 12 subjects in each condition. This mean z-transformed matrix was inversely transformed into the correlation matrix by a z-to-r transformation (Bewick et al., 2003). If correlation coefficients for the group matrix were directly averaged without z transformation, then those that were not normally

distributed would reflect only the central tendency of correlation values. To avoid these problems, I used r-to-z transformation and z-to-r reverse transformation for the group correlation matrices.

2.4.3 Brain functional connectivity

After computing the group correlation matrix, I computed the distance matrix by (1 - correlation value) between defined ROIs. The distance matrix for each condition was computed from an average z-to-r–transformed matrix. Then I applied graph filtration methods to the distance matrix with ε values taken discretely from the smallest distance to the largest distance between the ROIs for each condition. At the first iteration, ε was set to zero. By increasing ε, the edge between two regions was connected if the distance between them was smaller than ε. Logically, I defined the shape of the network by piecing together patches connected to nearest neighbor regions in an iterative fashion. If two regions were already connected directly or indirectly, I did not reconnect the two regions. When ε was larger than any distance between the two regions, the iteration was terminated since the graph shape could not change any further. This analysis allowed the shape of the brain network to be represented as both a MST and a SLD.

2.4.4 Networks using graph filtration in the persistent homology

The weighted functional network was usually thresholded at a particular value to interpret the functional connectivity between ROIs. In order to apply a conservative and proper threshold, a multiple comparison correction was performed over all edges for a given level of significance (Chen et al., 2008; Ferrarini et al., 2009; Salvador et al., 2005), or the sparsity of edges was controlled (Van den Heuvel et al., 2009; He et al., 2008). However, weaker correlations and connections tended to be ignored when edges were thresholded by significance level. This poses a problem because weak correlations may be particularly important for comparing networks between disease and normal groups (Bassett et al., 2012). In addition, because the sparsity of a network is defined as the ratio of the number of current edges to the number of all possible edges, it usually has to be controlled at maximum sparsity in order to compare networks between groups or conditions (Bassett et al., 2006). This results in the generation of redundant edges that can be difficult to interpret. Lee et al. (2011) proposed a new network modeling strategy that avoids this problem by comparing networks between groups and visualizing networks without any arbitrary thresholding in the persistent homology framework. Persistent homology is an approach that infers how local neighborhoods of an

object connect to each other to form a topology, from which graph filtration can be applied to assess whether or not any important features and structures persist (Edelsbrunner and Harer, 2008). Through use of this persistent homology framework, I was able to observe changes in topology over an entire range of thresholds rather than a single arbitrary threshold and identify the persistent features. To find the persistent features in the network, decomposing methods such as graph filtration were used in this study. During filtration, persistent topological features were observed, and the topological features were defined as an algebraic form known as the Betti number or connected component. The 0^{th} Betti number refers to the number of connected components in the network. In the graph filtration when the initial threshold was equal to 0, all nodes were disconnected and the number of connected components was equal to the number of nodes. By increasing the threshold, nodes were connected with each other. Graph filtration was continued by connecting nodes with correlation values larger than the specific threshold. Once all nodes were connected, graph filtration was stopped, and a single connected component remained. At that time, edges in all conditions were equal to 1 less than the number of nodes.

Graph filtration can be used to represent a sequence of finite numbers of nested, unweighted networks from the range of thresholds.

Using graph filtration, I found topologically invariant features such as the Betti number, which allowed for comparison between networks. During filtration, the number of connected components (or clustering) was calculated, and the change in the number of connected components was plotted as a barcode of the range of thresholds. This analysis represents global topological features in the network. However, it is difficult to identify geometric information about where changes between nodes occur from this data. If additional geometric information is contained within the barcode, it can be represented as a dendrogram. The dendrogram rearranges the bars in the barcode to form the SLD. The process of graph filtration is mathematically equivalent to constructing the single linkage hierarchical clustering and single linkage distance. In this step, if edges that change the number of connected components exist, it can be represented using MSTs. These techniques allowed the shape of the network to be defined. The MST is related to single linkage hierarchical clustering. In single linkage hierarchical clustering performed in this study, the distance between two clusters was computed as the distance between the closest elements in the two clusters. I represented the connected components in an algebraic form, known as the SLM. This method had been described in

Figure 2 and detailed in previous studies (Lee et al., 2011a, 2011b and 2011c).

2.4.5 Brain connectivity in whole brain regions

For whole brain connectivity, I used 90 cortical and subcortical regions (except for the regions in the cerebellum) defined by the AAL atlas as nodes (Table 1; Tzourio-Mazoyer et al., 2002). The distance matrix was computed as 1 - correlation using the correlation matrix between 90 regions. The distance matrix for each condition was computed from average z-to-r–transformed matrix resulting in a {90 × 90} distance matrix.

2.4.6 Task-related connectivity in activated regions

Speech-related functional connectivity was identified in the present study using a specific ROI-driven network in which nodes were based on task-related activations. For the task-related functional network, I first identified brain regions activated during the speech task. Speech conditions were as follows: audiovisual speech (AV), auditory

speech (A), and visual speech (V). Comparisons for speech-related regions were made between each speech condition and control condition, such as AV > C, A > C, and V > C. All comparisons generated from subject-level analysis were subsequently subjected to a group-level analysis in which the subject was considered as a random factor. A group analysis of a GLM with random effect was performed for each comparison using a one-sample t test. Comparison maps were then generated from these one-sample t tests ($p < 0.005$, uncorrected). From these results, the ROIs for nodes were identified where greater than 10 contiguous significant voxels were observed. If these identified regions overlapped, either the peak of the region with the higher t statistic was selected, or the region was divided into multiple regions. The regions were defined by placing a sphere with radius of 6 mm around the peak MNI coordinate of the activated region.

Mean BOLD signal values were extracted by averaging the activation within a sphere of 6–10 mm radius surrounding the maxima peak of activation, as has been established previously (He, 2011; Fox et al., 2009). For each speech condition network, the mean BOLD signal time series were extracted from each group. The time series values were taken from the first eigenvalue within the region and used as inputs for the task-related connectivity analysis. This analysis revealed

19 regions with significant task-related functional connectivity, as summarized in Table 3. These regions included 7 frontal regions, such as the inferior prefrontal gyrus (BA44), inferior prefrontal gyrus (BA45), middle frontal gyrus, premotor area, superior medial frontal gyrus, and ventral inferior frontal gyrus in the left hemisphere and the right medial frontal gyrus. The areas also included 7 temporal regions, including the anterior superior temporal gyrus and posterior superior temporal gyrus in the left hemisphere; the superior temporal gyrus, transverse temporal gyrus, and insula in the right hemisphere; and both the middle temporal gyri. Two subcortical regions (the cingulate gyrus and hippocampal gyrus in the left hemisphere) and 3 occipital regions (both the cuneus and left lingual gyrus) were also included. I constructed a {19 × 19} positive correlation matrix in each condition from which the distance matrix was computed. For the task-related functional connectivity analysis, I started graph filtration with ε values taken discretely from the smallest distance between regions to the largest distance between them in the distance matrix. The following procedure was the same as described in section 2.4.4.

2.4.7 Comparison of networks

For comparing the differences between networks in different conditions, I used a nonparametric permutation, in which the z-transformed metric was tested. I first compared the slopes of the barcodes between different conditions. Secondly, SLMs from each condition were converted into z values to ensure that the data was approximately normally distributed, using z = 1/2 log [(2 - r)/(r)], (i.e., 1-r was assigned instead of r from z = 1/2 log [(1 + r) / (1 - r)]). Then, these z-values were compared between conditions by z statistics using using $z = (Z_1 - Z_2)/ \sqrt{\left(\frac{1}{n1-3}\right) + \left(\frac{1}{n2-3}\right)}$. Differences in functional connectivity between different conditions were tested using a nonparametric permutation test with 5000 repetitions (He et al., 2008). Nonparametric permutation testing was used to test the likelihood that the observed difference between different conditions occurred by chance (the null hypothesis).

Measurement of whole brain connectivity used a vector per measurement, since the network was built per condition and per subject. The nonparametric test was performed by reassigning the condition measurement for each subject. Each subject's data (including the slope

of the barcode and each entry of single linkage distance) was randomly sampled to generate a pseudo-condition. The difference between the pseudo-conditions was tested using an unpaired t-test. After repeating the procedure 5000 times, the t-statistics served as a null distribution to find significance at $p < 0.05$ (two-tailed).

For measurement of task-specific connectivity, a nonparametric test was performed by resampling the subject data, since the data included a scalar per condition. To compare networks between two conditions, all subjects of the conditions were pooled together and then randomly reassigned to one condition, yielding two pseudo-condition data matrices. On each pseudo-data matrix, the slope of the barcode and single linkage distance were estimated. This procedure was repeated 5000 times to generate the null distribution. Significance was defined as $p < 0.05$ (two-tailed).

3. Results

3.1 Behavioral performance

The behavioral results show that the correct response performance was greatest during the AV condition (74.6%), followed by the A (57.2%) and V (31%) conditions ($F [2, 32] = 18.16$, $p < 0.0001$). Since subjects with normal hearing lack the ability of lip-reading, the V condition was the most difficult condition. Most subjects inferred the sentence in the V condition on the basis of prior language knowledge. A post hoc test (Scheffe's t-test) indicated that the difference in performance between the V and either AV ($P < 0.0001$) or A ($P < 0.005$) conditions was statistically significant.

3.2 Task-related functional connectivity

3.2.1 Results of GLM analysis

Figure 5 shows the activation maps observed during the AV, A, and V conditions relative to the control condition at uncorrected $p < 0.005$ (with extent threshold size = 10). During the AV condition, significant activation was observed in both the superior temporal gyri/sulci (STS/STS). The A condition was characterized by activation

of both the STG/STS, the left middle temporal gyrus, both cuneus, lingual gyrus, left hippocampal gyrus, and left dorsolateral/ventral inferior prefrontal regions. In particular, a greater extent of significant activation was found in the left anterior to posterior superior temporal regions during the A condition. During the V condition, significant activation was found in the left dorsolateral/ventral inferior prefrontal regions, left superior/middle frontal region, right insula, right medial frontal, and right cingulate regions (Figure 5, Table 2).

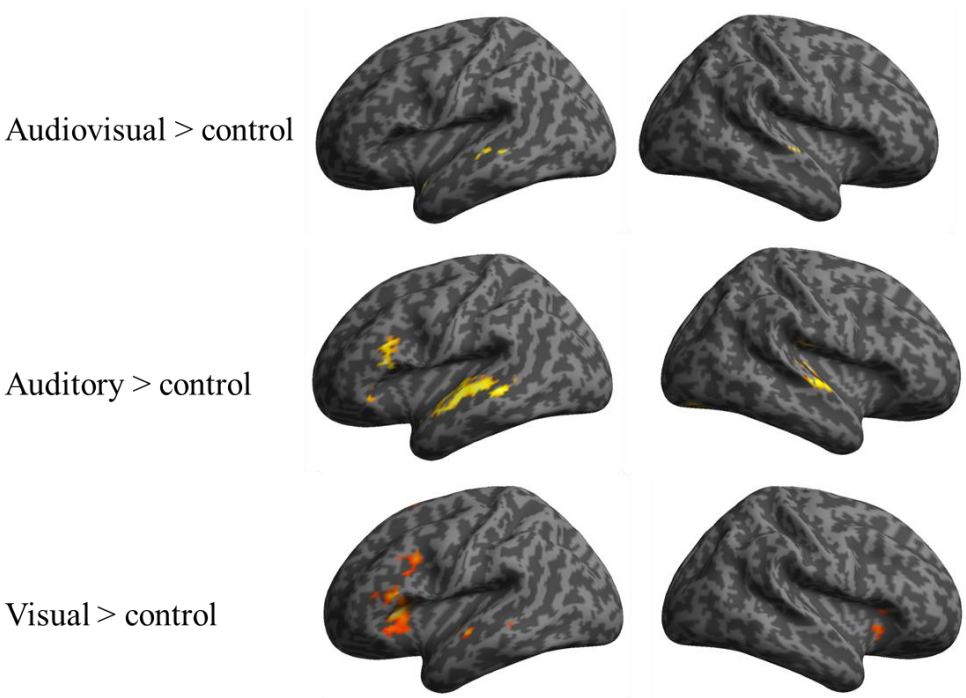


Figure 5. Areas with significantly increased activation, displayed through whole brain analysis for each condition compared to control. Brain regions in which BOLD signal changes were observed are shown in red-yellow color. (uncorrected $p < 0.005$, extend threshold $k > 10$).

Table 2. The significant brain regions during speech condition relative to control condition for functional connectivity

Condition		Region	BA	Talairach coordinates			t score	cluster size
				x	y	z		
AV	L	Anterior superior temporal G	38	-40	7	-19	5.93	179
	L	Middle temporal G	21	-51	-31	0	3.59	39
	R	Superior temporal G	21	61	-10	-1	3.52	11
A	L	Inferior frontal G	45	-51	22	12	5.45	234
	L	Ventral inferior frontal G	47	-46	27	-6	3.93	64
	L	Middle temporal G	21	-50	-12	-6	5.27	924
	L	Anterior superior temporal G	38	-44	7	-14	4.08	19
	R	Middle temporal G	21	65	-28	-4	3.64	14
	R	Transverse temporal G	41	46	-21	10	5.06	274
	R	Superior temporal G	22	61	-12	1	4.45	
	L	Cuneus	30	-24	-71	11	4.89	117
	R	Cuneus	18	10	-85	8	4.45	82
	L	Lingual G	18	-14	-78	-6	4.09	
	L	Cingulate G	23	-10	-16	28	3.75	10
	L	Hippocampal G		-28	-24	-4	5.07	48
V	L	Inferior frontal G	45	-34	24	4	7.48	732
	L	Inferior frontal G	44	-38	9	29	4.37	205
	L	Middle frontal G	9	-53	19	34	4.28	
	L	Superior medial frontal G	8	-10	14	53	3.82	53
	L	Middle frontal G or premotor	6	-44	4	50	3.54	28
	R	Medial frontal G	9	8	33	32	4.18	132
	L	Middle temporal G	21	-51	-14	-4	5.1	63
	L	Anterior superior temporal G	38	-44	7	-12	4.66	68
	L	Posterior superior temporal G		-57	-37	6	3.6	38
	R	insula	13	34	20	3	4.53	137
	L	Cingulate G	23	-6	-14	27	4.52	75

G:gyrus, Talairach coordinates for local maxima of clusters which were composed of more than ten contiguous significant voxels (uncorrected $p<0.005$)

3.2.2 Common functional connectivity within speech conditions

Nineteen ROIs were selected to investigate task-related functional connectivity (Table 3, Figure 6 (d)). The mean BOLD signals were extracted from the ROIs within a 6 mm radius of local maxima. Contrast images of speech condition compared to control condition from subjects were used for speech-related connectivity.

Figure 6 (a) shows the positive correlation matrices of the AV, A, and V conditions. Figure 6 (b) shows the correlation maps with thresholding at uncorrected $p < 0.01$. In the AV condition, positive correlation in 26 out of 171 connections were statistically significant, while 7 and 27 positive connections were identified in the A and V conditions, respectively. Many significantly positive connections were identified in the frontal regions in the AV and V conditions. Such relationships included connections between the inferior frontal region (BA44) and middle frontal (BA9) or ventral inferior frontal regions (BA47), between the inferior frontal region (BA44) and posterior superior temporal region (BA22), and between the ventral inferior frontal (BA47) and middle frontal (BA9) or posterior superior temporal regions (BA22) in the left hemisphere. Connections were also observed between the right superior temporal (BA22) and the lingual gyrus (BA18).

In the AV and A conditions, there were common connections between left premotor (BA6) and left ventral inferior frontal (BA47) or right medial frontal regions (BA9), between the superior medial frontal region (BA8) and cuneus (BA30/18), between ventral inferior frontal (BA47) and middle temporal (BA21) areas, and between the occipital regions in left hemisphere. In the A and V conditions, there was only one significant connection between the left middle temporal region (BA21) and right superior temporal gyrus (BA22).

SLMs of speech conditions are displayed in Figure 6 (c). In the AV condition, fronto-temporal connectivity was stronger than other connections, with the exception of connections with the anterior superior temporal region. In addition, other regions' connections with both right temporal and subcortical regions were relatively weak. In SLM of the A condition, however, the pattern of connected structures was generally not region-specific. Most connections of the A condition connected later than connections in the other conditions. In the V condition, connections of other regions with the right transverse temporal region were the last to be connected.

Table 3. Brain regions that were selected for task-specific functional connectivity

ROI order		region	BA	Talairach coordinates			t score	cluster size
				x	y	z		
Frontal regions								
1	L	Inferior frontal G	45	-34	24	4	7.48	732
2	L	Inferior frontal G	44	-38	9	29	4.37	205
3	L	Middle frontal G	9	-53	19	34	4.28	
4	L	premotor	6	-44	4	50	3.54	28
5	L	Superior medial frontal G	8	-10	14	53	3.82	53
6	L	Ventral inferior frontal G	47	-46	27	-6	3.93	64
7	R	Medial frontal G	9	8	33	32	4.18	132
Temporal regions								
8	L	Anterior superior temporal G	38	-40	7	-19	5.93	179
9	L	Middle temporal G	21	-50	-12	-6	5.27	924
10	L	Posterior superior temporal G		-57	-37	6	3.6	38
11	R	Middle temporal G	21	65	-28	-4	3.64	14
12	R	Superior temporal G	22	61	-12	1	4.45	
13	R	Transverse temporal G	41	46	-21	10	5.06	274
14	R	insula	13	34	20	3	4.53	137
Limbic regions								
15	L	Cingulate G	23	-6	-14	27	4.52	75
16	L	Hippocampal G		-28	-24	-4	5.07	48
Occipital regions								
17	R	Cuneus	18	10	-85	8	4.45	82
18	L	Cuneus	30	-24	-71	11	4.89	117
19	L	Lingual G	18	-14	-78	-6	4.09	

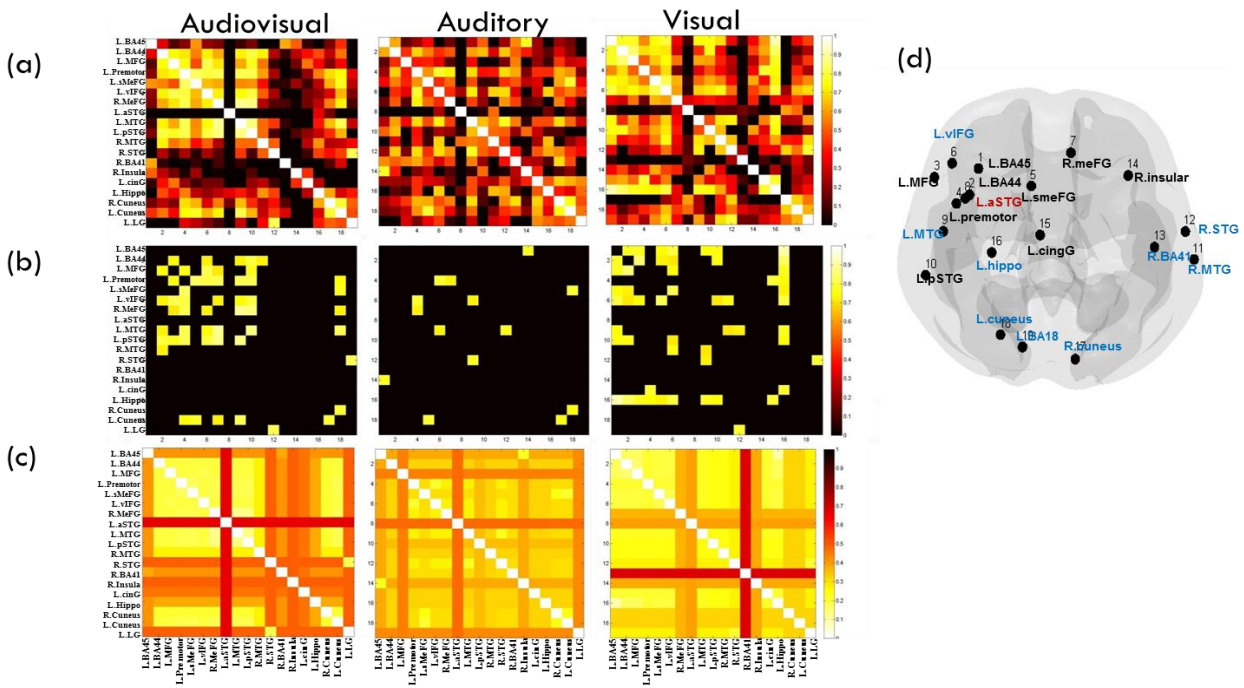


Figure 6. The connectivity matrix. (a) Each connectivity map represents the positive correlation matrix in each condition. (b) Uncorrected connectivity maps ($p < 0.01$). (c) Each connectivity map represents the single linkage matrix. The heatmap indicates shorter distance between nodes, with yellow indicating higher correlation. (d) Brain regions were selected from GLM analysis for functional connectivity. All brain regions are described in Table 3.

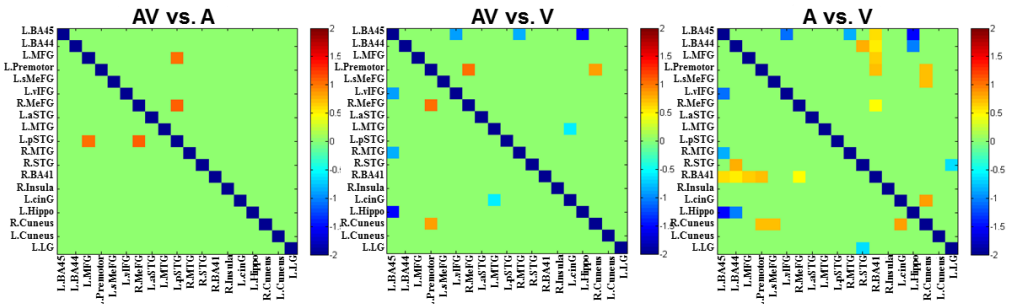
3.2.3 Differences in functional connectivity between different speech conditions

Figure 7 shows the observed differences in functional connectivity between the three speech conditions. Compared to the A condition, the AV condition showed significantly stronger correlations between the left posterior superior temporal region and the left and right medial frontal regions (left panel of Figure 7 (a)). Compared to the V condition, the AV condition showed stronger correlations between the left premotor and both the right medial frontal area and the right cuneus (middle of Figure 7 (a)). Compared to both the AV and A conditions, the V condition elicited significantly stronger correlations between the left inferior frontal region (BA45) and either the left ventral frontal, right middle frontal, or left hippocampal regions. Additionally, compared to the AV condition, the V condition had stronger correlations between the middle temporal and cingulate regions in the left hemisphere. Compared to the A condition, the V condition showed stronger correlations between the inferior frontal and hippocampal region in the left hemisphere and between the right superior temporal and occipital regions (right panel of Figure 7 (a)). Compared to A condition, the A condition elicited significantly stronger correlations between the transverse temporal region and

several frontal regions, and between the premotor region and cuneus.

Nonparametric permutation testing was used for comparing the SLMs (Figure 7 (b)). Compared to the both AV and A conditions, the V condition showed a stronger connection between the inferior frontal and hippocampal region in the left hemisphere. Compared to AV condition, the V condition showed stronger connections between the middle temporal and cingulate regions in the left hemisphere, while compared to the V condition, the A condition showed stronger connections between the right transverse temporal and other regions. There was no significant difference between the AV and A conditions.

(a) Correlation matrix



(b) Single linkage matrix

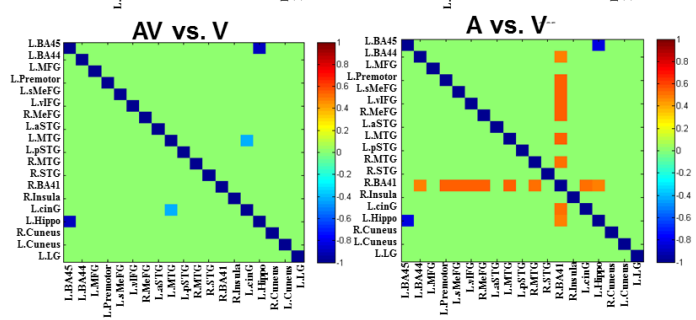


Figure 7. Differences in the functional connectivity between different speech conditions. (a) Connectivity difference maps were computed between correlation matrices using nonparametric permutation testing at $p < 0.05$. (b) Differences between single linkage matrices (SLMs) were also computed using nonparametric permutation testing at $p < 0.05$.

3.2.4 The shape of networks in speech conditions

Figure 8 (a) shows the slopes of connected components across the entire range of filtration values during the three speech conditions. Differences in the shape of task-related functional brain networks were estimated by comparing the barcode slopes (i.e., beta values of regression) using nonparametric permutation testing. The slopes of barcodes of the A condition and both the AV and V conditions were significantly different (Figure 8 (b)).

In Figure 9 (a), the barcode and dendrogram of each condition is displayed, and Figure 9 (b) represents the connected component matrix at each filtration value. In the dendrogram plotted in Figure 9 (a), the maximum filtration thresholds for connected components were 0.67, 0.47, and 0.68 for AV, A, and V conditions, respectively. Despite differences in the shape of the barcode slopes shown in Figure 8 (a), interpretation of specific connected components on the SLD is problematic. To avoid this issue, several filtration values were selected to compare the different patterns of connected components. The selected filtration values were 0.15, 0.20, 0.25, 0.35, 0.40, 0.50, and 0.70. Filtration values were selected heuristically using the SLD. The generated clusters at each filtration value are summarized in Table A1.

Figure 9 (b) shows the sequences of connected components in

the AV condition in the order they were identified, where the first clustering of connected components was between the left posterior superior temporal region and middle temporal region, and then the clustering was extended within frontal regions and between frontal and occipital regions. Unlike the AV condition, the V condition showed the first clustering of connected components between the hippocampal and inferior frontal regions in the left hemisphere, and connections were later extended within the frontal regions. The last connected component (as expected) was the primary auditory cortex. The pattern of clustering of connected components in the A condition was global and lacked region specificity.

Figure 10 shows the MSTs by the single linkage distance for task-related connectivity. The number of edges in each MST was 18. As shown in Figure 10 (a), the only common edge between all three MSTs was between both the cuneus.

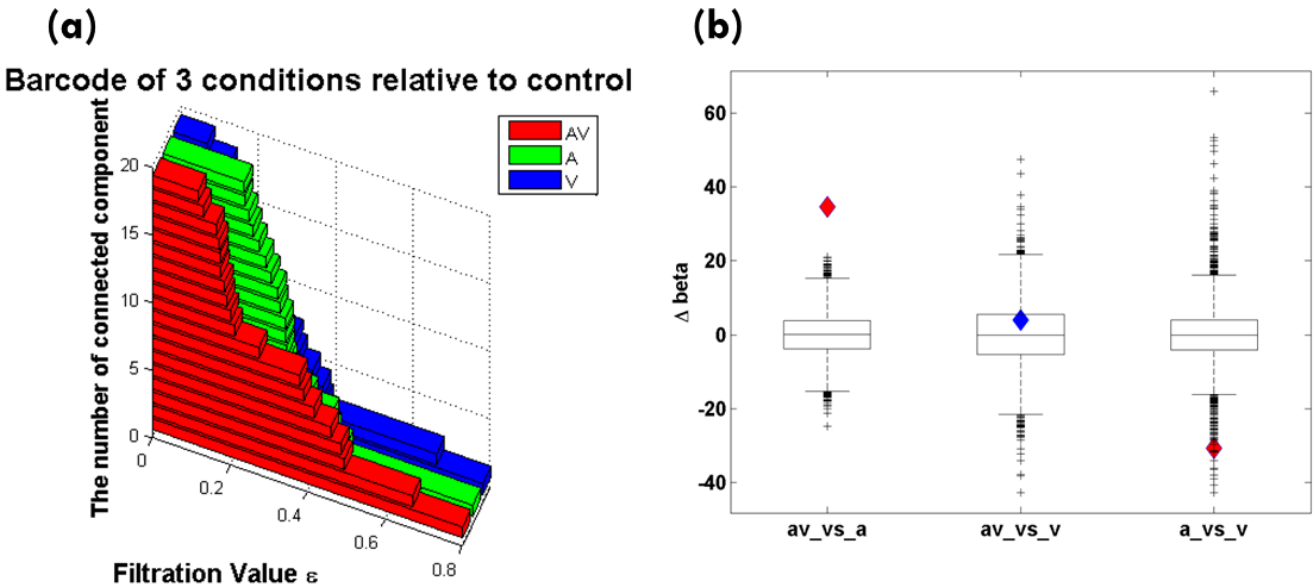


Figure 8. Barcodes representing functionally connected components during the three speech conditions across a range of filtration values, ε. (a) The vertical axis displays the number of connected component and the horizontal axis shows filtration value. Red barcodes represent the AV condition, green for A, and blue for V conditions. (b) Differences in the slope of barcodes between conditions display. For differences of functional connectivity through filtration values, nonparametric permutation testing was performed and significance was defined as $p < 0.05$. AV, A and V represent audiovisual, auditory, and visual conditions, respectively.

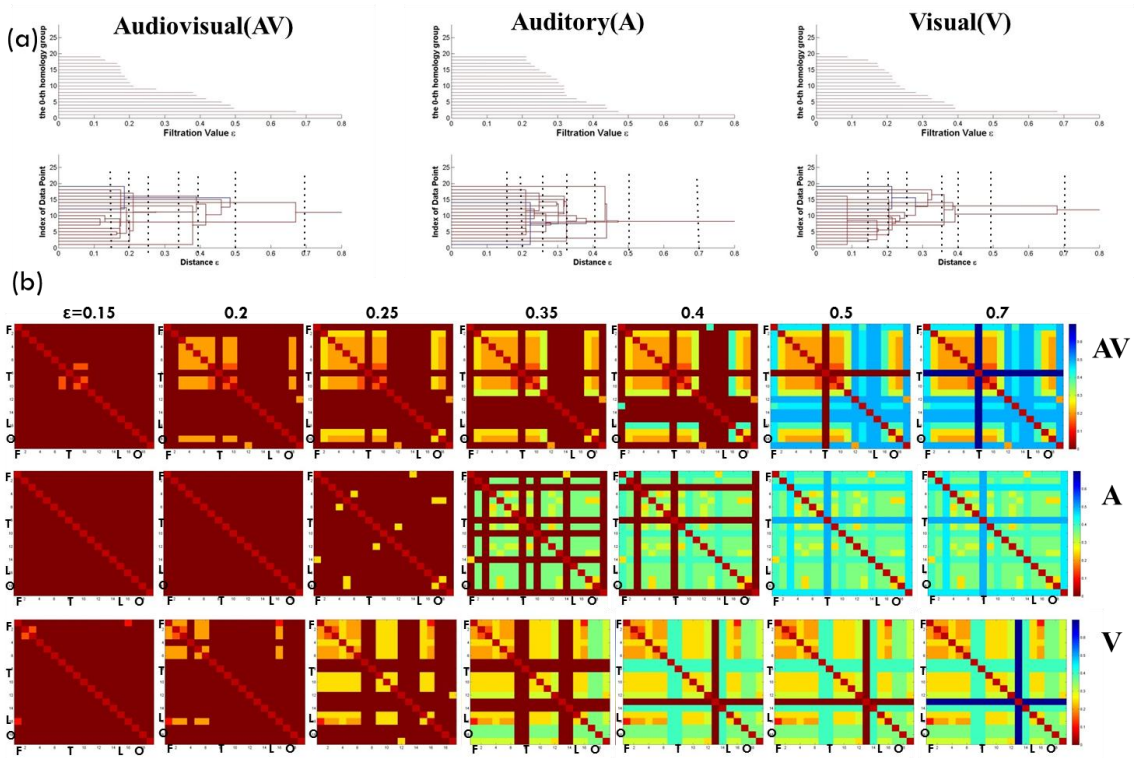


Figure 9. The connected component matrix through filtration values. (a) Lines of connected components and single linkage dendrogram of speech conditions (b) Connected component matrix through filtration values of 0.15, 0.20, 0.25, 0.35, 0.40, 0.50, and 0.70.

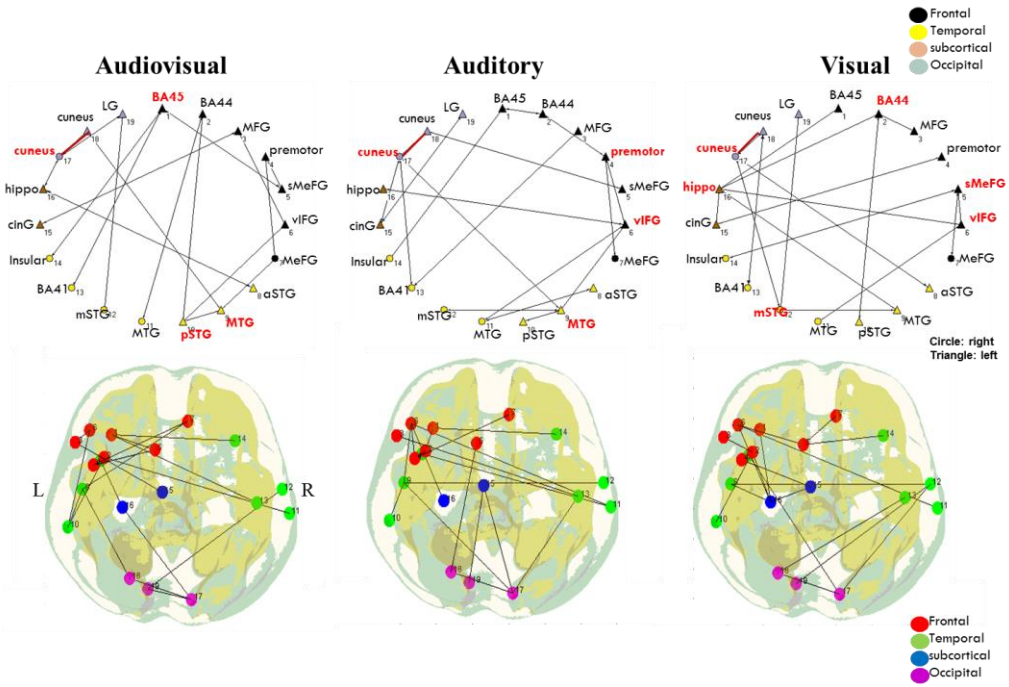


Figure 10. The minimum spanning trees (MSTs) of three speech conditions. (a) The MSTs formed a circular pattern. Color displays brain regions, such as black, yellow, brown, and light blue for frontal, temporal, subcortical, and occipital regions, respectively. (b) The MSTs were illustrated in the corresponding brain structure. Node colors were as follows: red, green, blue, and magenta for frontal, temporal, subcortical, and occipital regions, respectively. Common edge was displayed as a red line on the circular pattern.

3.3 Whole brain functional connectivity

3.3.1 Functional connectivity within conditions

I examined the positive correlations between all pairs of 90 nodes for the four conditions (AV, A, V, and C). The positive correlation matrices of the four conditions are shown in Figure 11 (a). In Figure 11 (a), the Pearson correlation matrix of each condition is displayed without any threshold. Each correlation matrix for the weighted network showed their average correlation strengths as 0.53, 0.50, 0.47, and 0.46 for AV, A, V, and C conditions, respectively. In Figure 11 (b), each correlation matrix was thresholded at the same value (correlation coefficients, $r > 0.7$). The number of connections at a given threshold in each condition was 436, 285, 241, and 205 for AV, A, V, and C conditions, respectively. Inter- and intra-hemispheric connections (i.e., bright rectangles along the 2 diagonal lines) were more prominent than other connections across all conditions.

SLMs generated using graph filtration are shown in Figure 11 (c). As shown in Figure 11 (c), connections between the occipital regions tended to be strong and emerge earlier across all conditions. The colorbar in Figure 11 (c) represents filtration values, in which bright colors (smaller distance values) represents higher correlations.

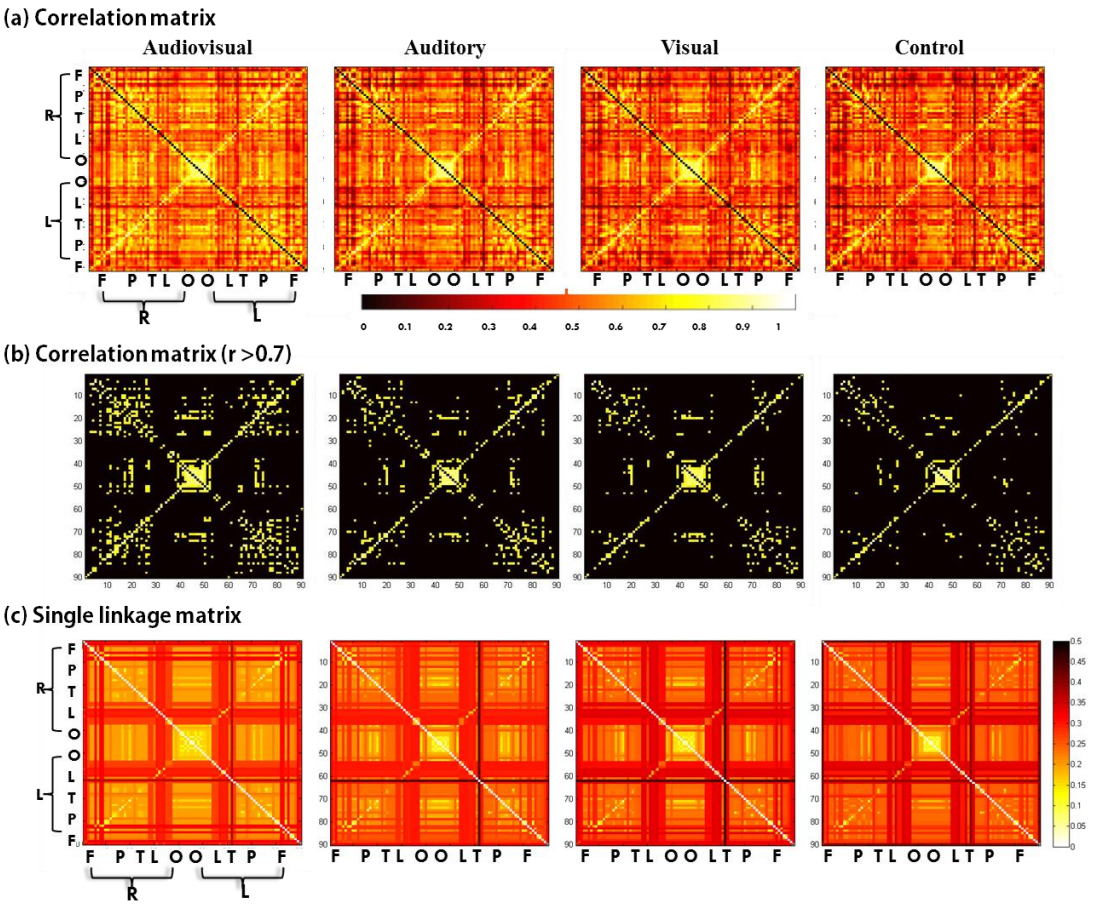

(a) Correlation matrix
Audiovisual
Auditory
Visual
Control
R
L
F P T L O O L T P F
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
(b) Correlation matrix (r >0.7)
(c) Single linkage matrix

Figure 11. Distance and single linkage matrices. In the top panel (a), the distance matrices were obtained by 1 - correlation after estimation between 90 regions of interests. In the middle panel (b), each correlation matrix was thresholded at an arbitrary value (correlation coefficient > 0.7). Correlations in two diagonal lines represent correlations between homologous regions in the bilateral hemisphere. In the bottom panel (c), the single linkage matrices were obtained based on the distance matrix. The heatmap denotes sequence of component merging during analysis, with brighter colors representing earlier merging. Yellow color represents higher correlation. Correlations along the two diagonal lines represent correlations both between regions within the same hemisphere and between homologue regions in the bilateral hemisphere.

3.3.2 Differences in functional connectivity between different conditions

For differences in correlation matrices between conditions, nonparametric permutation testing was performed. Six pairs were used for comparison. Differences between the speech conditions (i.e., AV, A, and V) and control condition were as follows: stronger correlation was observed in the AV versus C conditions between the right inferior frontal triangular part (R.F3T) and bilateral supplementary motor area (SMA), between the middle frontal gyrus (R.F2) and dorsolateral superior frontal gyrus (R.F1), and between the medial superior frontal gyrus (R.F1M) and inferior parietal region (R.P2) in the right hemisphere. Stronger connections in the A condition compared to the C condition were observed between the right amygdala (R.AMYG) and left thalamus (L.THA), and there were no significantly different connections between the V and C conditions at uncorrected probability of $p < 0.001$.

Differences in connections between the three speech conditions were as follows: the connection between the left middle temporal pole (L.T2P) and left middle frontal orbital part (L.F2O) was stronger in the AV condition compared to the A condition. The connections between the right inferior occipital gyrus (R.O3) and left precentral gyrus

(L.PRE) and between the left thalamus (L.THA) and left olfactory cortex (L.OC) were stronger in the AV condition than in the V condition. Connections between both the middle frontal gyri (F2) and between the putamen and middle temporal pole in the left hemisphere were stronger in the A condition than in the V conditions at an uncorrected probability of $p < 0.01$ (Supplementary Figure S1).

A difference in the SLMs of the AV and C conditions was observed between the right middle frontal and right dorsolateral superior frontal regions. Between the A and V conditions, a significant difference was observed between the middle frontal and superior frontal gyri in the right hemisphere. No significant differences at a given threshold were observed for any other pairs.

3.3.3 Shape of networks for whole brain connectivity

Functional connectivity was assessed for every possible threshold for all four conditions using graph filtration. Figure 12 shows connected components in the functional network. These topological changes were represented by barcode of the 0^{th} Betti number. As the filtration value (i.e., threshold) increased, the number of connected components decreased. In Figure 12 (a), the vertical axis represents the number of connected components, and the horizontal axis shows the

filtration value. As the filtration value approached zero, the number of connected components approached the number of nodes (90). Maximum filtration thresholds, (i.e., the level at which all nodes became 1 connected component) were 0.36, 0.44, 0.46, and 0.40 for the AV, A, V, and C conditions, respectively. To assess differences in the connected components, the slopes of barcodes of all conditions were compared using nonparametric permutation testing at $p < 0.05$. Significant difference in the slope was observed only between the AV and V conditions.

Figure 12 (b) shows the barcodes rearranged according to node index where the vertical axis represents the index of nodes. This plot represents the single linkage distance, which shows the order of local connections. However, interpretation of the connected components using this analysis was overly complicated. Therefore, the structure of connected components was observed by selecting several different filtration values (Supplementary Figure S2). The filtration values were 0.10, 0.16, 0.20, 0.22, 0.28, and 0.30 for each condition. Figure S2 shows the connectivity pattern across this range of filtration values. When the filtration value was sufficiently large, most nodes were connected to each other, while filtration value near zero resulted in most nodes being disconnected.

The SLD can also be represented by MSTs (Supplementary Figure S3). The number of edges of all MSTs was 89. Edges in the MSTs represent edges that made changes of connected structures. Figure S3 (a) shows MSTs in the circle with left-right symmetry. Figure S3 (b) shows MSTs in the corresponding brain structures. Figure S3 (c) shows MSTs with common edges across all conditions. Among the 89 edges of the MST, 55 were common. Common edges mainly consisted of edges between homologous regions in both hemispheres, between occipital regions, and between frontal regions (except for a few edges). Each MST without common edges in MSTs is displayed in Supplementary Figure S4.

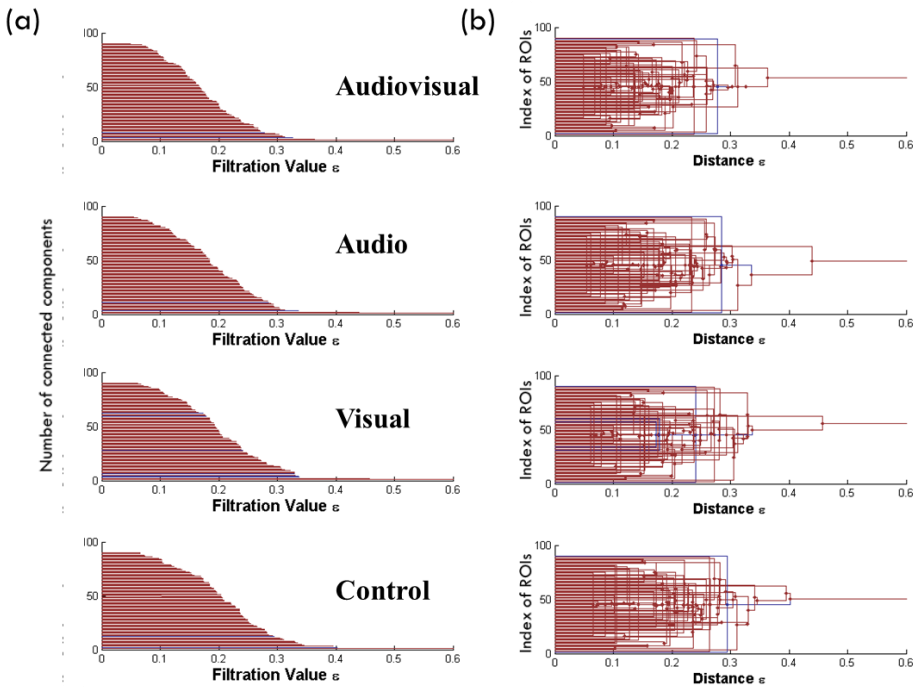


Figure 12. Connected components of brain network revealed using graph filtration. (a) Barcodes representing connected components through varying filtration values during each condition. (b) Single linkage dendrograms of each condition represented by barcodes rearranged according to node index.

4. Discussion

This study aimed at investigating the functional brain networks involved in audiovisual speech processing using graph filtration in a persistent homology framework. In particular, I investigated the mechanisms that coordinate communication and integration of audiovisual speech in the network of task-relevant brain structures. Subtle network differences in audiovisual speech integration and auditory or visual speech processing were revealed in this study. This study revealed three main findings: (i) audiovisual speech processing was associated with strong connectivity of the left posterior superior temporal region and also involved relationships with frontal regions, such as fronto-temporal and fronto-occipital connections. By contrast, auditory speech processing relied more on connectivity with the superior temporal region. (ii) Visual speech processing revealed remarkable connectivity between the hippocampus and inferior frontal region; this suggests that visual speech processing may be related to speech production memory. (iii) Compared to the AV and V conditions, the A condition did not show specific local connectivity, but instead revealed global connectivity between all regions. Additionally, in

whole brain networks, compared to other conditions, the AV condition was associated with stronger overall correlation strengths. All three speech conditions showed connectivity between bilateral homologous regions. Elucidation of connectivity patterns via the graph filtration methods in persistent homology performed in this study could give important insight into the perceptual stages involved in audiovisual speech integration.

4.1 Brain networks during speech tasks

Human speech has auditory and visual properties. Although auditory speech can be used to communicate in the absence of visual speech, visual speech improves speech communication in noisy environments. Because the information coming from the auditory and visual modalities is independent, combining information from both modalities (multimodal or multisensory processing) allows for more accurate perception. Sumby and Pollack (1954) showed that speech perception was enhanced in noisy environments when congruent visual speech was delivered simultaneously. The influence of visual speech on auditory speech perception was also reported by Grant and Seitz (2000). They showed that correlated information from the face with mouth movement improved detection of noisy auditory messages.

The neural representations of audiovisual integration have been investigated using many imaging tools, including fMRI, PET, EEG, and MEG. Evidence from such studies suggests that the left STS/STG is a critical area for multisensory integration. Further, the associative auditory cortex including the left posterior superior temporal sulcus is believed to be the primary neural substrate of audiovisual speech integration. Conversely, others have argued that in addition to the temporal regions, the left inferior frontal cortex is a crucial brain area for speech perception (Hickok and Poeppel, 2000). In this study, I investigated how the observed functional connectivity between regions depended on which modality was used to convey speech. I hypothesized that differences in speech modality would manifest as differences in functional connectivity. The results reveal striking differences in functional connectivity between different speech modalities. Although the functional neuroanatomy involved in audiovisual speech perception was difficult to characterize precisely, task-dependent functional connectivity was observed between activated regions. The left superior/middle temporal region mainly showed functional connectivity with auditory or visual regions conveying auditory or visual speech, respectively.

4.1.1 Regions activated during speech conditions

In many neuroimaging studies, results can be biased because of the inconsistent criteria used to classify multisensory integration. However, Love et al. (2011) recently proposed three criteria to identify multisensory integration. First, responses must be supra-additive, meaning that the bimodal response must be greater than the sum of both unimodal responses. Second, the bimodal response must be greater than the largest unimodal response. Last, the bimodal response must be greater than the mean of the unimodal responses. Calvert et al. (2000) reported that supra-additive responses were observed in the temporal, occipital, parietal, and frontal regions and were focused in the left posterior superior temporal sulcus. Szycik et al. (2008) reported that both the superior temporal cortices were the regions of cross-modal integration. In the present study, the activation pattern evoked by the speech condition compared to the control condition is shown in Figure 5. As in our previous study (Kang et al., 2006), activation was observed in auditory speech-related temporal regions such as both the superior or middle temporal regions (AV > C and A > C). Both the temporal regions were activated to a lesser extent in the AV condition than the A speech condition, probably because audiovisual speech processing is more automatic and natural. No significant voxels were found outside

of temporal regions in the AV condition. During the A condition, additional activation was observed in the left dorsolateral inferior prefrontal regions (BA44/BA45). These regions were activated to mimic sequences of mouth movements given a particular auditory cue.

In all speech conditions, linguistic information was conveyed with the auditory or visual modality, as was non-linguistic information (such as simple mouth movements or noise). The linguistic information contained within visual speech is often referred to as lip-reading or silent speech. Lip-reading–evoked activation has been observed in the STG/STS (Calvert et al., 1997; Puce et al., 1998). In our study, we observed sizeable activations in the left inferior prefrontal and ventral frontal regions (BA44 and BA47) during the V condition. The left Broca's area (BA44) was activated by hand or mouth movements as well as speech representation (Rizzolatti et al., 2004).

Rizzolatti et al. (2004) reported that neurons responding to the observation of actions were present not only in the inferior frontal region, but also in the left STS. Left STS/STG is a critical area responding to multimodal integration. However, Nath et al. (2011) reported that left STS/STG showed inter-individual variability in integration capacity of audiovisual speech perception at the behavioral level. The left anterior ventral inferior prefrontal region (BA47) in

particular, might be involved during guessing of ambiguous words from unrecognizable sentences. In addition, the medial prefrontal region (BA8/10) was activated during the V condition. This activation in the V condition might also be included in guessing the speaker's mouth movement to identify the speaker's intention.

4.1.2 Functional connectivity during audiovisual speech processing

The activation results reported in this study are consistent with other evidence suggesting that the left STS/STG is a critical brain area for auditory and visual multisensory integration (Beauchamp, 2005b). Wright et al. (2003) proposed the left posterior superior temporal cortex as the key structure for integrating meaningful multisensory information (Wright et al., 2003). Particularly, the left posterior superior temporal region has been hypothesized to be an important cortical region for the integration of auditory and visual speech information (Calvert et al., 2000; Okada et al., 2009). That region was also reported to be activated by congruent auditory and visual stimuli (Berstein et al., 2008; Miller and D'Esposito, 2005; Sekiyama et al., 2003; Campbell et al., 2001). This region is known to be the human homologue of the superior temporal polysensory area in macaques, which receives projections from auditory and visual association

cortices (Lewis and Van Essen, 2000). During audiovisual speech perception, auditory cortex processes audio information, visual cortex processes visual speech information, and superior temporal regions integrate bimodal information (Calvert et al., 2000; Olson et al., 2002; Macaluso et al., 2003; Sekiyama et al., 2003; Wright et al., 2003; Hickok and Poeppel, 2007; Bernstein et al., 2008; Campbell, 2008).

From the SLMs shown in Figure 9, the functional connection involved in the AV condition was formed between left posterior superior/middle temporal regions and medial frontal regions early from the sequence of filtration. The medial frontal regions are anatomically linked with the superior temporal regions. Connectivity between medial frontal and superior/middle temporal regions has been associated with attentional state such as task solving or grasping speaker's intention (Kang et al., 2006). This finding suggests that audiovisual speech integration emerges in fronto-temporal circuitry as an interactive process with auditory and visual speech. Lee et al. (2011) suggested that left fronto-temporal circuitry was determined by both the temporal structure of auditory signals as well as the amount of cross-modal practice that enabled articulatory representations. They also suggested that fronto-temporal connectivity might integrate auditory and visual information into a speech percept guided by prior articulatory gestural

representations (Lee et al., 2011). As filtration values increased, the clustering pattern of connectivity in audiovisual speech integration expanded to include connections between the frontal regions and occipital and temporal regions. That is, functional connectivity of audiovisual speech integration expanded by connecting with frontal regions.

There is evidence that auditory and visual speech cues interact with each other, especially in noisy environments where speech perception is enhanced when auditory and visual speech information is congruent (Ma et al., 2009; Ross et al., 2007a, b). The functional connectivity results suggest that the connections mediating this phenomenon might be between the left pSTG and Broca's area. Visual speech processing showed this connectivity when visual speech contributed meaningful information. In addition to connectivity with frontal regions, there was stronger connectivity between occipital and temporal regions during the AV condition than during the A or V conditions (Love et al., 2011). The occipito-temporal connectivity can be interpreted as forming relatively early in audiovisual speech perception, potentially having a mechanistic role in multisensory integration (Joassin et al., 2011). The SLM in Figure 9 shows that the connection strengths between visual or auditory regions and left

temporal multisensory areas are increased during the AV condition compared to the A or V condition. This finding supports previous results demonstrating greater functional coupling between sensory cortices and multisensory cortex (such as superior temporal cortex) when audiovisual stimuli were temporally congruent (Kreifelts et al., 2007; Noesselt et al., 2007) Hemispheric differences were explored in this study. Connections with the right superior temporal regions (ROIs 12 and 13) formed later in the AV condition than in the A condition. Kang et al. reported that audiovisual speech processing showed left-lateralized areas of activation (Kang et al., 2006). The right hemisphere seems to play a smaller role in speech integration or perception (Wolmetz et al., 2010). Although activation was observed bilaterally in the superior temporal regions during speech processing (Hickok and Poeppel, 2007), the connectivity of the right temporal brain regions may not be essential when there is sufficient speech information. In my results, however, the right brain regions were not identified as ROIs as often as brain regions in the left hemisphere, a result that could be related to the effectiveness of left-lateralized speech processing.

4.1.3 Functional connectivity during auditory speech processing

During the A condition, functional connectivity was formed relatively slowly than during the other conditions (middle column of Figure 9 (b)). The audiovisual connectivity pattern revealed global functional connectivity without connectivity between specific local regions. Particularly, there was connectivity among the temporal regions in the right hemisphere. The A condition in this study also included visual information, although the visual information contained meaningless mouth movements rather than speech information. This required that subjects pay more attention to auditory speech information. Functional connections between the right insular and left inferior frontal regions were also observed. These regions are known to participate in phonological mediation for lip-reading in both lexical and non-lexical processing (Paulesu et al., 2003). One possibility is that the lack of local functional connectivity in the A condition reflects the suppression of unnecessary visual information in the early connectivity. Although the activation of inhibitory circuits was not directly evaluated in this study, it is possible that an intrinsic inhibition mechanism caused the auditory speech processing connectivity to be formed relatively later between specific regions.

Unexpectedly, connections with the anterior temporal regions were the last to be identified in audiovisual and auditory speech tasks. A previous study reported that the anterior temporal region responded to higher-order linguistic information and intelligible speech stimuli. Visser et al. (2010) reported that the anterior temporal region might underlie an amodal semantic system, the specific pattern of which is influenced by stimulus type according to variations in the degree of connectivity with modality specific areas. Thus, although the role of the anterior temporal regions remains unclear, it may have been identified last due to a lack of direct functional connections between the anterior temporal and other regions. Another possible explanation is that this effect arose due to the methodological limitation that all regions are composed of one connected component for a functional network.

4.1.4 Functional connectivity during visual speech processing

It is generally agreed that the ventral stream in auditory speech processing supports the perception and recognition of auditory speech (Hickok and Poeppel, 2000, 2004 and 2007). The dorsal stream in auditory speech maps sensory or phonological representations onto articulatory motor representations (Hickok and Poeppel, 2007). Normal subjects find it difficult to recognize visual speech. One strategy to help

facilitate understanding of visual speech is to mimic the speaker's articulation as well as to understand the meaning of the sentence.

In the V condition, the most informative functional connectivity was between the left inferior frontal and hippocampal regions. The connected component between these areas was identified early on as a strong connection. In addition to this study, the study by Kang et al. (2006) reported activation of the left frontal regions in the V condition. Visual speech generated robust activation in the inferior frontal gyrus (BA44 and 45) and generally evoked greater activation on the left than the right side (Calvert and Campbell, 2003). The left hippocampal region was activated often in another cross-modal study of face and voice integration (Love et al., 2011). Functional connectivity between the left inferior frontal and hippocampal regions might be interpreted as using previous individual knowledge about speech production in the early parts of visual speech processing. In this study, the V condition was composed of a visual speech stimulus and auditory noise. Lip-reading requires involvement of language regions such as the frontal regions as well as visual regions (Paulesu et al., 1997). Additionally, Hall et al. (2005) reported that visual speech engaged the middle temporal gyrus, inferior and middle frontal gyri, and inferior parietal regions in the left hemisphere, all of which are included in the lexical

stages of spoken language. The MST in Figure 10 shows connectivity between the right middle superior temporal region and lingual region in the V condition as well as in the AV condition but not the A condition; this might represent the effect of meaningful visual speech. Functional connectivity between the left pSTG and inferior frontal gyrus (BA44) was common in both the AV and V conditions. The left posterior superior temporal regions play a significant role in natural visual speech, including audiovisual speech binding and integration in hearing people. Therefore, visual speech processing and audiovisual speech processing were likely to share common mechanisms of visual speech analysis.

For connections with left STS/STG, visual speech did not always involve sound-based coding of visual input. Left STG/STS activity in visual speech processing was strongly related to visual speech skill (Hall et al, 2005) or the phonological stage of visual speech processing, but was not related to simple mouth movements. Nevertheless, the finding that visual speech elicits left STG activation has been replicated in many studies (Calvert and Campbell, 2003; MacSweeney et al., 2002). Figure 9 (a) shows a SLD for the visual condition in which the pattern of clustering was first formed between frontal regions and between occipital regions, and then later between

temporal regions. As expected, connections with primary auditory cortex involved in auditory signal processing emerged later. Such sequences of merged connectivity do not necessarily imply a temporal flow of connectivity, but might indicate the functional relationships of brain regions. From this result, the visual speech processing in normal subjects represented meaningful speech-related actions. It has been proposed that visual speech processing might share a common representation framework in the perception of meaningful action (Decety and Grezes, 1999).

4.2 Whole brain functional networks

This study used graph filtration methods to find whole brain networks of task-related functional connectivity in a threshold-independent manner. Whole brain functional connectivity revealed strong connections in both inter-hemispheric areas (such as between bilaterally symmetric regions) and intra-hemispheric areas (such as anatomically adjacent regions).

The inter-hemispheric pattern of connectivity contained symmetric regions of each hemisphere. This finding is consistent with anatomical studies demonstrating that fibers of the corpus callosum link symmetric regions in the brain (Zaidel and Iacoboni, 2003). In particular, inter-

hemispheric connectivity was observed between occipital regions more readily than other regions across all experimental conditions. This result seemed to be related to the fact that the stimuli used in this study were moving faces. Within each condition, I found that most connections started locally between anatomically adjacent regions. Connectivity strengths in the AV condition were larger than in the other conditions. This result is consistent with a previous study in which correlations tended to diminish when subjects exerted more effort in cognitive tasks or unnatural environments (Ginestet et al., 2011a).

Context-independent functional connectivity was observed in occipital regions, although the specific connectivity strengths varied by condition. I also identified strong connections between other lobes during auditory and visual speech. In this study of fMRI imaging, task-related stimuli were presented together amidst continuous noise produced by the fMRI itself. In this situation, visual speech was important to understand speech sentences as well as auditory speech. Therefore, during audiovisual speech processing, connectivity between many brain regions showed stronger connectivity than that observed in the visual speech task. In fact, this study used face stimuli with naturally speaking mouth movement or meaningless mouth movement. The difference of AV and A conditions was in visual mouth movement.

Congruent visual speech information with auditory speech remarkably increased correlation strengths between brain regions related to speech production.

The ventral stream in auditory speech processing supports the perception and recognition of auditory speech whereas dorsal stream in auditory speech maps sensory or phonological representations onto articulatory motor representations (Hickok and Poeppel, 2007). In the visual speech condition of this study in particular, it was too difficult to recognize visual speech in the absence of auditory cues, so mimicking the speaker's articulation was a useful strategy to understand the sentence.

In the whole brain connectivity (Supplementary Figure S3 (c)), MSTs from all conditions showed 55 common connections. These results also showed that tasks requiring significant effort and those with distracting non-task-related information, (i.e., visual speech and non-speech information) displayed pervasive reduction in overall functional connectivity. This finding suggests that patterns of whole brain functional connectivity are similar during speech tasks, although we did not investigate the functional connectivity of other cognitive tasks or functional connectivity of the resting state. That is, whole brain connectivity appears to be context-independent and to represent the

intrinsic brain state during the completion of cognitive tasks rather than functional connectivity related to a specific cognitive function.

4.3 Methodological contributions and limitations

In this framework, the network structure can be changed depending on a range of specified thresholds, thereby avoiding the pitfalls of arbitrary threshold selection. Comparison of functional connectivity between tasks at given threshold can identify correlation relationships, while this method also allows for the shape of networks to be investigated over a continuous range of thresholds via the graph filtration method. The correlation measures used in this study were transformed into distance measures from which the single linkage distance measure could be computed. These procedures can be visualized as 0^{th} Betti number and barcodes that represent the shape of a network. The MSTs of the audiovisual and auditory speech conditions are shown in Figure 10. Alexander-Bloch et al. (2010) presented the MST as an alternative thresholding method instead of global thresholding to construct functional brain networks. Although the structure of the MST itself is biologically implausible because by definition it contains no loops or triangles, the MST still can be used for identifying the minimal connected skeleton of a brain network. In addition, basic MSTs can be modified by adding extra edges according to a local thresholding rule. In this study, the MST was constructed by single linkage distance with the smallest distance between clusters.

This topology might not represent differences of true speech-specific functional networks. Future work may overcome this limitation of biological plausibility by adding holes (i.e., 1st Betti number) to the MST (Lee et al., 2012).

5. Conclusions

Functional connectivity between posterior superior temporal regions and frontal regions is remarkable during audiovisual speech integration. Connectivity between the temporal regions and auditory or visual areas seems to be driven by auditory or visual speech modality. The mechanism by which the brain incorporates visual speech information to improve auditory speech perception seems to rely on connectivity with motor areas related to speech production rather than visual sensory areas. This study also provides evidence that functional connectivity elucidated by graph filtration is different from speech processing sequences, depending on speech modality.

REFERENCES

Achard, S., Salvador, R., Whitcher, B., Suckling, J., Bullmore, E. (2006) A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. J Neurosci 26, 63-72.

Alexander-Bloch AF, Gogtay N, Meunier D, Birn R, Clasen L, Lalonde F, Lenroot R, Giedd J, Bullmore ET. (2010) Disrupted modularity and local connectivity of brain functional networks in childhood-onset schizophrenia. Front Syst Neurosci. 8;4:147.

Adler, R. J., Bobrowski, O., Borman, M. S., Subag, E., Weinberger, S. (2010) Persistent homology for random fields and complexes. ArXiv e-prints.

Bassett DS, Bullmore ET. (2009) Human brain networks in health and disease. Curr Opin Neurol; 22(4):340-347.

Bassett DS, Nelson BG, Mueller BA, Camchong J, and Lim KO. (2012) Altered resting state complexity in schizophrenia. NeuroImage 59, 2196–2207.

Bernstein LE, Auer ET Jr, Wagner M, Ponton CW. (2008) Spatiotemporal dynamics of audiovisual speech processing. Neuroimage 39 (1) : 423-435.

Bewick, V., Cheek, L., Ball, J. (2003). Statistics review 7: Correlation

and regression. Critical Care 7:451-459.

Buccino G, Binkofski F, Fink GR, Fadiga L, Fogassi L, Gallese V, Seitz RJ, Zilles K, Rizzolatti G, Freund HJ. (2001) Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. Eur J Neurosci. 13(2):400-404.

Buccino G, Lui F, Canessa N, Patteri I, Lagravinese G, et al. (2004a). Neural circuits involved in the recognition of actions performed by non-conspecifics: an fMRI study. J. Cogn. Neurosci. 16:1–14.

Buckner, R.L., Sepulcre, J., Talukdar, T., Krienen, F.M., Liu, H., Hedden, T., Andrews-Hanna, J.R., Sperling, R.A., Johnson, K.A. (2009) Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease. J Neurosci. 29, 1860-73.

Beauchamp MS. (2005) See me, hear me, touch me: multisensory integration in lateral occipital-temporal cortex. Curr Opin Neurobiol. 15: 145–153.

Bullmore, E. and Sporns, O. (2009) Complex brain networks: graph theoretical analysis of structural and functional systems. Nat Rev Neurosci 10 186–198.

Caeyenberghs K, Leemans A, Heitger MH, Leunissen I, Dhollander T,

Sunaert S, Dupont P, Swinnen SP. (2012) Graph analysis of functional brain networks for cognitive control of action in traumatic brain injury. Brain135: 1293-307.

Callan DE, Callan AM, Kroos C, Vatikiotis-Bateson E. (2001) Multimodal contribution to speech perception revealed by independent component analysis: a single-sweep EEG case study. Brain Res Cogn Brain Res. 10 (3): 349-53.

Calvert, G.A., Bullmore, E.T., Brammer, M.J., Campbell, R., Williams, S.C., McGuire, P.K., Woodruff, P.W., Iversen, S.D., David, A.S. (1997) Activation of auditory cortex during silent lipreading. Science 276, 593– 596.

Calvert GA, Campbell R, Brammer MJ. (2000) Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. Curr Biol. 10 (11): 649-657.

Calvert GA and Campbell R. (2003) Reading speech from still and moving faces: The Neural substrates of visible SpeechEvidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. Curr Biol. 10 (11): 649-657.

Campbell, R. (1998) How brains see speech: the cortical localization of speechreading in hearing people. In: Campbell, R., Dodd, B.,

Burnham, D. (Eds.), Hearing by Eye II: Advances in Psychology of Speechreading and Auditory-Visual Speech. Psychology Press, Hover, pp. 177–193.

Campbell, R., MacSweeney, M., Surguladze, S., Calvert, G., McGuire, P., Suckling, J., Brammer, M.J., David, A.S. (2001) Cortical substrates for the perception of face actions: an fMRI study of the specificity of activation for seen speech and for meaningless lower-face acts (gurning). Brain Res. Cogn. Brain Res. 12, 233–243.

Chen, Z. J., He, Y., Rosa-Neto, P., Germann, J., Evans, A. C. (2008) Revealing modular architecture of human brain structural networks by using cortical thickness from MRI. Cereb. Cortex 18, 2374–2381.

Cordes D, Haughton VM, Arfanakis K, Carew JD, Turski PA, Moritz CH, Quigley MA, Meyerand ME. (2000) Frequencies contributing to functional connectivity in the cerebral cortex in "resting-state" data. AJNR Am J Neuroradiol, 22(7), 1326-33.

Damoiseaux JS, Rombouts SA, Barkhof F, Scheltens P, Stam CJ, Smith SM, Beckmann CF. (2006) Consistent resting-state networks across healthy subjects. Proc Natl Acad Sci 103(37):13848-53

Edelsbrunner, H. and Harer, J. (2008) Persistent homology - a survey. Contemporary Mathematics 453, 257–282.

Ferrarini, L., Veer, I. M., Baerends, E., van Tol, M.-J., Renken, R. J., van der Wee, N. J., Veltman, D. J., Aleman, A., Zitman, F. G., Penninx, B. W., van Buchem, M. A., Reiber, J. H., Rombouts, S. A., Milles, J., (2009) Hierarchical functional modularity in the resting-state human brain. Hum. Brain Mapp. 30, 2220–2231.

Fletcher, P.C., Happe, F., Frith, U., Baker, S.C., Dolan, R.J., Frackowiak, R.S., Frith, C.D., (1995) Other minds in the brain: a functional imaging study of ‘‘theory of mind’’ in story comprehension. Cognition NeuroImage 32, 423–431.

Fox, M.D., Zhang, D., Snyder, A.Z., Raichle, M.E. (2009) The global signal and observed anticorrelated resting state brain networks. J. Neurophysiol. 101, 3270–3283.

Fornito A, Yoon J, Zalesky A, Bullmore ET, Carter CS. (2011) General and specific functional connectivity disturbances in first-episode schizophrenia during cognitive control performance. Biol Psychiatry 70, 64-72.

Fransson P, Marrelec G. (2008) The precuneus/posterior cingulate cortex plays a pivotal role in the default mode network: Evidence from a partial correlation network analysis

Neuroimage.42(3):1178-1184.

Frackowiak, R., Friston, K., Frith, C., Dolan, R., Price, C., Zeki, S., Ashburner, J., Penny, W. (2004) Human brain function. Academic Press, San Diego. (Eds.)

Friederici, A.D., Meyer, M., von Cramon, D.Y. (2000) Auditory language comprehension: an event-related fMRI study on the processing of syntactic and lexical information. Brain Lang. 74, 289–300.

Friederici, A.D., Ruschemeyer, S.A., Hahne, A., Fiebach, C.J. (2003) The role of left inferior frontal and superior temporal cortex in sentence comprehension: localizing syntactic and semantic processes. Cereb. Cortex 13, 170– 177.

Friston, K.J. (1994). Functional and effective connectivity in neuroimaging: Human Brain Mapping, 22-78

Ginestet, CE., Simmons, A. (2011a) Statistical parametric network analysis of functional connectivity dynamics during a working memory task. Neuroimage 55, 688-704.

Grady C. L., McIntosh AR and Craik F. I.M. (2003) Age-Related Differences in the Functional Connectivity of the Hippocampus During Memory Encoding. HIPPOCAMPUS 13:572–586.

Grant KW. And Seitz PF. (2000) The use of visible speech cues for

improving auditory detection of spoken sentences. J Acoust Soc Am. 108:1197-208.

Greicius MD, Krasnow B, Reiss AL, et al. (2003) Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. Proc Natl Acad Sci ;100:253–258

Greicius MD, Supekar K, Menon V, Dougherty RF. (2009) Resting-state functional connectivity reflects structural connectivity in the default mode network. Cereb Cortex. 19(1):72-78

Hall, DA., Fussell C., Summerfield, Q. (2005) Reading fluent speech from talking faces: typical brain networks and individual differences. Journal of cognitive neuroscience, 17, 939-953.

He, Y., Chen, Z., Evans, A. (2008) Structural insights into aberrant topological patterns of large-scale cortical networks in Alzheimer's disease. J Neurosci 28, 4756-4766.

He, Y., Chen, Z., Gong, G., Evans, A. (2009) Neuronal networks in Alzheimer's disease. Neuroscientist. 15, 333-50.

He, Y. and Evans, A. (2010) Graph theoretical modeling of brain connectivity. Current Opinion in Neurology, 23:341–350.

He H, Liu TT. (2012) A geometric view of global signal confounds in resting-state functional MRI. Neuroimage 59:2339-2348.

Hickok, G. and Poeppel, D. (2000) Towards a functional

neuroanatomy of speech perception. Trends in Cognitive Sciences 4. No. 4.

Hickok, G. and Poeppel, D. (2004) Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. Cognition 92 67–99.

Hickok, G. and Poeppel, D. (2007) The cortical organization of speech Processing Nature reviews| neuroscience Vol 8.

Joassin F, Pesenti M, Maurage P, Verreckt E, Bruyer R, Campanella S. (2011) Cross-modal interactions between human faces and voices involved in person recognition. Cortex. ;47(3):367-376.

Kang, E, Lee DS, Kang H, Hwang CH, Oh SH, Kim CS, Chung JK, Lee MC. (2006) The neural correlates of cross-modal interaction in speech perception during a semantic decision task on sentences: a PET study. Neuroimage. 32(1):423-431.

Kreifelts B, Ethofer T, Grodd W, Erb M, Wildgruber D. (2007) Audiovisual integration of emotional signals in voice and face: an event-related fMRI study. Neuroimage. 37(4):1445-1456.

Lee L, Friston K, Horwitz B. (2006) Large-scale neural models and dynamic causal modelling. Neuroimage 30: 1243–1254.

Lee, H., Chung, M. K., Kang, H., Kim, B. N., Lee, D. S. (2011a) Computing the shape of brain network using graph filtration and

gromov-haudorff metric. In: 14th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI). Vol. 6891. Toronto, Canada, pp. 289–296.

Lee, H., Chung, M. K., Kang, H., Kim, B. N., Lee, D. S. (2011b) Discriminative persistent homology of brain networks. In: IEEE International Symposium on Biomedical Imaging (ISBI). Chicago, IL, pp. 841–844.

Lee, H., Chung, M. K., , Kang, H., Kim, B. N., Lee, D. S. (2011c) Persistent network homology from the perspective of dendrograms. In: 17th Annual Meeting of the Organization for Human Brain Mapping (HBM). Quebec City, Canada.

Lee H, Noppeney U. (2011) Physical and Perceptual Factors Shape the Neural Mechanisms That Integrate Audiovisual Signals in Speech Comprehension. The Journal of Neuroscience, 31(31):11338 –11350.

Lewis J.W, Van Essen D.C. (2000) Corticocortical connections of visual, sensorymotor, and multimodal processing areas in the parietal lobe of the macaque monkey. The Journal of comparative neurology, 428:112 –137.

Lord LD, Allen P, Expert P, Howes O, Lambiotte R, McGuire P, Bose SK, Hyde S, Turkheimer FE. (2011) Characterization of the

anterior cingulate's role in the at-risk mental state using graph theory. Neuroimage. 56(3):1531-1539.

Love SA, Pollick FE, Latinus M. (2011) Cerebral correlates and statistical criteria of cross-modal face and voice integration. Seeing Perceiving. 24(4):351-367.

Ma WJ, Zhou X, Ross LA, Foxe JJ, Parra LC. (2009) Lip-reading aids word recognition most in moderate noise: a Bayesian explanation using high-dimensional feature space. PLoS One. 4(3):e4638.

Macaluso E, George N, Dolan R, Spence C, Driver J. (2004) Spatial and temporal factors during processing of audiovisual speech: a PET study. Neuroimage. 21(2):725-732.

McGurk, H., MacDonald, J. (1976). Hearing lips and seeing voices. Nature 264, 746-748.

McIntosh AR, Lobaugh NJ. (2004) Partial least squares analysis of neuroimaging data: applications and advances. Neuroimage. 23 Suppl 1:S250-263.

Miller LM, D'Esposito M. 2. (2005) Perceptual fusion and stimulus coincidence in the cross-modal integration of speech J Neurosci 25(25):5884-5893.

Molinari E, Baraldi P, Campanella M, Duzzi D, Nocetti L, Pagnoni G,

Porro CA. (2012) Human Parietofrontal Networks Related to Action Observation Detected at Rest Cereb Cortex.

Morris J. S., Ohman A. and Dolan R. J. (1999) A subcortical pathway to the right amygdala mediating ‘‘unseen’’ fear. Proc. Natl. Acad. Sci. USA Vol. 96, pp. 1680–1685.

Nakamura T, Hillary FG, Biswal BB. (2009) Resting network plasticity following brain injury. PLoS One14;4(12):e8220.

Nath AR, Beauchamp MS. (2011) Dynamic changes in superior temporal sulcus connectivity during perception of noisy audiovisual speech. J Neurosci. 2;31(5):1704-1714.

Noesselt T, Rieger JW, Schoenfeld MA, Kanowski M, Hinrichs H, Heinze HJ, Driver J. (2007) Audiovisual temporal correspondence modulates human multisensory superior temporal sulcus plus primary sensory cortices. J Neurosci. 27(42):11431-1141.

Okada K, Hickok G. (2009) Two cortical mechanisms support the integration of visual and auditory speech: a hypothesis and preliminary data. Neurosci Lett. 20;452(3):219-223.

Olson IR, Gatenby JC, Gore JC. (2002) A comparison of bound and unbound audio-visual information processing in the human cerebral cortex. Brain Res Cogn Brain Res. 14(1):129-138.

Pachou E, Vourkas M, Simos P, Smit D, Stam CJ, Tsirka V, Micheloyannis S.(2008) Working memory in schizophrenia: an EEG study using power spectrum and coherence analysis to estimate cortical activation and network behavior. Brain Topogr. 21(2):128-37.

Pantazatos SP, Talati A, Pavlidis P, Hirsch J. (2012) Decoding unattended fearful faces with whole-brain correlations: an approach to identify condition-dependent large-scale functional connectivity. PLoS Comput Biol. 8(3):e1002441.

Paulesu E, Perani D, Blasi V, Silani G, Borghese NA, De Giovanni U, Sensolo S, Fazio F. (2003) A functional-anatomical model for lipreading. J Neurophysiol. 90:2005-2013.

Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, Vogel AC, Laumann TO, Miezin FM, Schlaggar BL, Petersen SE. (2011) Functional network organization of the human brain. Neuron 72, 665-678.

Puce A, Allison T, Bentin S, Gore JC, McCarthy G. (1998) Temporal cortex activation in humans viewing eye and mouth movements. J Neurosci.18(6):2188-99.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A. & Shulman, G. L. (2001) Proc. Natl. Acad. Sci.

USA 98, 676–682.

Ramsey JD, Hanson SJ, Glymour C. (2011) Multi-subject search correctly identifies causal connections and most causal directions in the DCM models of the Smith et al. simulation study. Neuroimage 58, 838-848.

Reijneveld JC, Ponten SC, Berendse HW, Stam CJ. (2007) The application of graph theoretical analysis to complex networks in the brain. Clin Neurophysiol18 (11):2317-2331.

Rizzolatti1, G and Craighero, L. (2004) THEMIRROR-NEURON SYSTEM Annu. Rev. Neurosci.. 27:169–192.

Ross LA, Saint-Amour D, Leavitt VM, Javitt DC, Foxe JJ. (2007a) Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. Cereb Cortex. 17(5):1147-1153.

Ross LA, Saint-Amour D, Leavitt VM, Molholm S, Javitt DC, Foxe JJ. (2007b) Impaired multisensory processing in schizophrenia: deficits in the visual enhancement of speech comprehension under noisy environmental conditions. Schizophr Res 97(1-3):173-183.

Salvador, R., Suckling, J., Coleman, M.R., Pickard, J.D., Menon, D., Bullmore, E. (2005) Neurophysiological architecture of

functional magnetic resonance images of human brain. Cereb Cortex. 15, 1332-1342.

Salvador R, Martínez A, Pomarol-Clotet E, Gomar J, Vila F, Sarró S, Capdevila A, Bullmore E. (2008) A simple view of the brain through a frequency-specific functional connectivity measure. NeuroImage 39 279–289.

Sekiyama K, Kanno I, Miura S, Sugita Y. (2003) Auditory-visual speech perception examined by fMRI and PET. Neurosci Res.47(3):277-287.

Smith SM, Fox PT, Miller KL, Glahn DC, Fox PM, Mackay CE, Filippini N, Watkins KE, Toro R, Laird AR, Beckmann CF. (2009) Correspondence of the brain's functional architecture during activation and rest. PNAS 106 13040.

Smith SM, Miller KL, Salimi-Khorshidi G, Webster M, Beckmann CF, Nichols TE, Ramsey JD, Woolrich MW. (2011) Network modelling methods for FMRI. Neuroimage 54, 875-891.

Smith SM. (2012). The future of fMRI connectivity. Neuroimage 5x, xxx-xxx.

Sporns, O., Zwi, J.D. (2004) The small world of the cerebral cortex. Neuroinformatics 2, 145-162.

Sporns, O., Tononi, G., Kotter, R. (2005) The human connectome: A

structural description of the human brain. PLoS Comput Biol. 1, e42.

Sporns, O. (2011) The human connectome: a complex network. Ann N Y Acad Sci.

Stam CJ, Reijneveld JC. (2007) Graph theoretical analysis of complex networks in the brain. Nonlinear Biomed Phys. 1 5;1(1):3.

Stam, C.J (2010) Characterization of anatomical and functional connectivity in the brain: A complex networks perspective. International Journal of Psychophysiology 77 186–194.

Sumby W.H, Pollack I. Visual contribution to speech intelligibility in noise. J Acoust Soc Am. 1954;26:212–215.

Supekar, K., Menon, V., Rubin, D., Musen, M., Greicius, M.D. (2008) Network analysis of intrinsic functional brain connectivity in Alzheimer's disease. PLoS Comput Biol. 4, e1000100.

Szycik GR, Tausche P, Münte TF. (2008) A novel approach to study audiovisual integration in speech perception: localizer fMRI and sparse sampling. Brain Res. 18;1220:142-149.

Szycik GR, Jansma H, Münte TF. (2009) Audiovisual integration during speech comprehension: an fMRI study comparing ROI-based and whole brain analyses. Hum Brain Mapp. 30(7):1990-9.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M. (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage. 15, 273-289.

Van den Heuvel, M. P., Stam, C. J., Kahn, R. S., Hulshoff Pol, H. E. (2009) Efficiency of functional brain networks and intellectual performance. J. Neurosci. 29, 7619–7624.

Varoquaux G, Baronnet F, Kleinschmidt A, Fillard P, Thirion B. (2010) Detection of brain functional-connectivity difference in post-stroke patients using group-level covariance modeling. Medical Image Computing and Computer Assisted Intervention Society, vol. 13, 200-208.

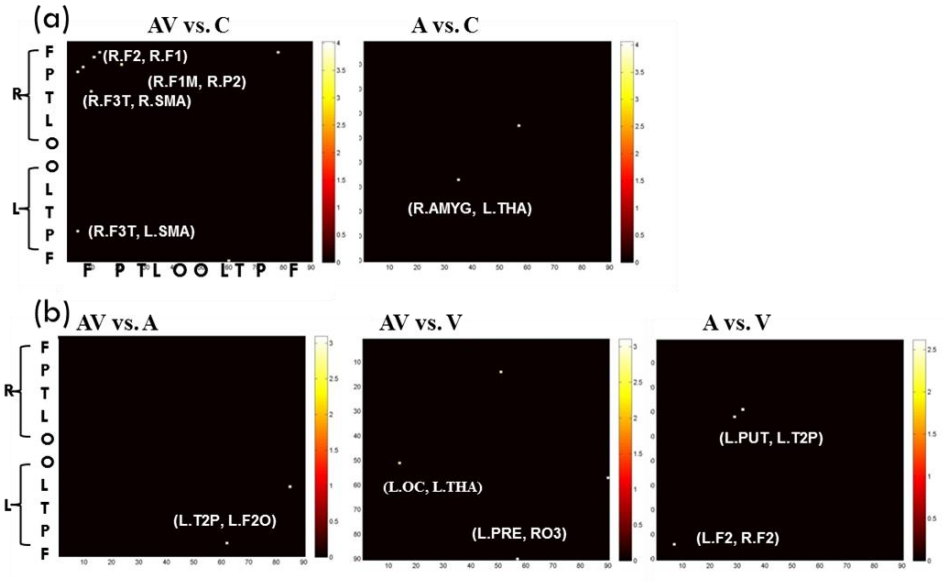
Visser M, Jefferies E, Lambon Ralph MA. (2010) Semantic processing in the anterior temporal lobes: a meta-analysis of the functional neuroimaging literature. J Cogn Neurosci. 22(6):1083-1094.

Wolmetz M, Poeppel D, Rapp B. (2010) What does the right hemisphere know about phoneme categories? J Cogn Neurosci. 23(3):552-569.

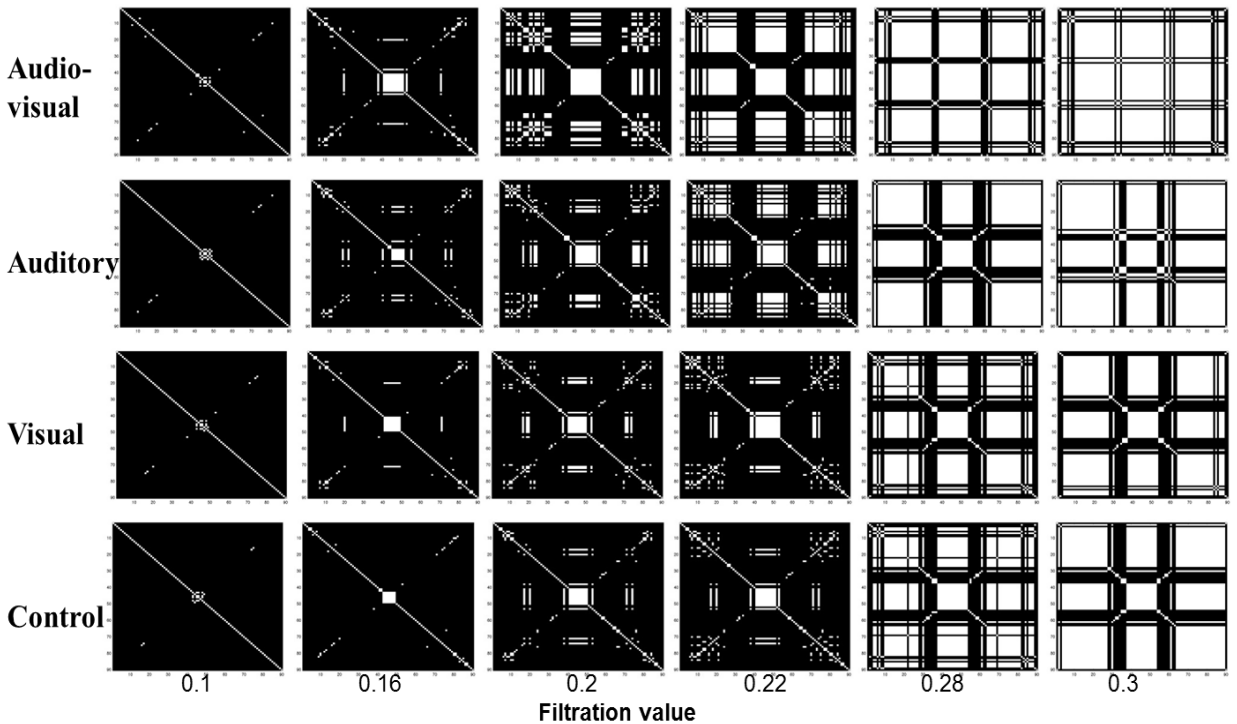
Wright TM, Pelphrey KA, Allison T, McKeown MJ, McCarthy G.

(2003) Polysensory interactions along lateral temporal regions evoked by audiovisual speech. Cereb Cortex. 13(10):1034-1043.

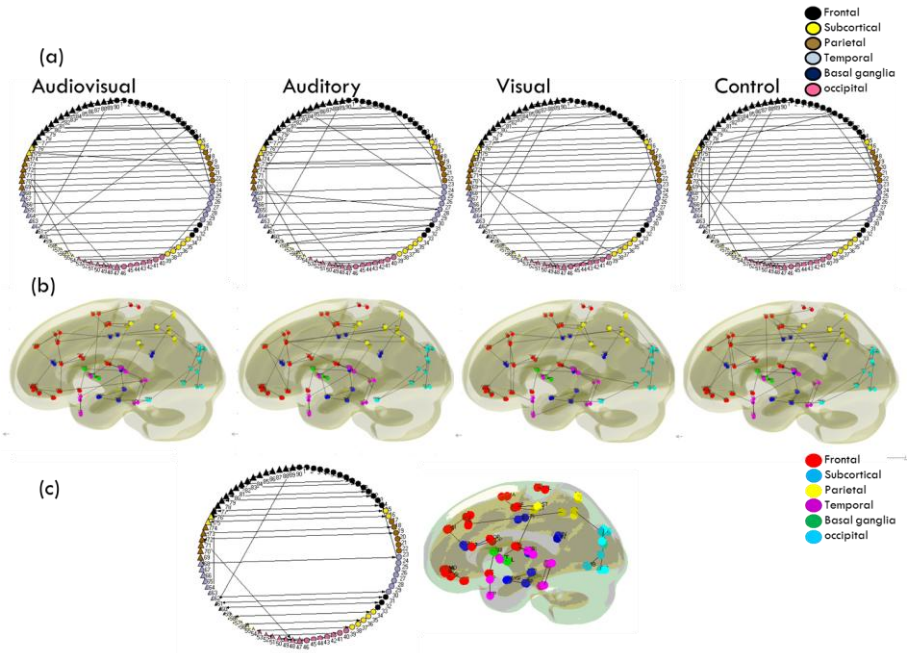
Zaidel, E., Iacoboni, M. (2003) The Parallel Brain: The Cognitive Neuroscience of the Corpus Callosum. MIT Press, New York.



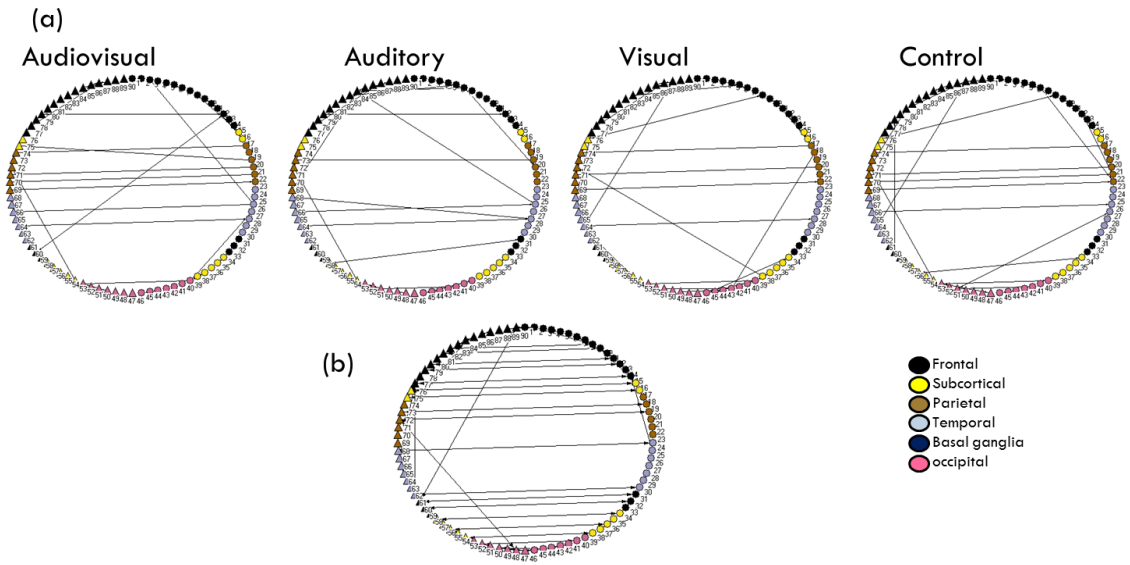
Supplementary Figure S1. Differences in correlation matrices between conditions. (a) Connectivity differences were found by using nonparametric permutation testing between audiovisual vs. control and auditory vs. control at uncorrected $p < 0.001$ (b) The signifi cance in three comparisons: audiovisual vs. auditory, audiovisual vs. visual, and auditory vs. visual at uncorrected $p < 0.01$. White rectangle denotes statistically significant connection.



Supplementary Figure S2. The connected component matrix through filtration values. Each row represents filtration values, 0.10, 0.16, 0.20, 0.22, 0.28, and 0.30 respectively. Each column represents the audiovisual (AV), auditory (A), visual (V), and control (C) conditions, respectively.



Supplementary Figure S3. The minimum spanning trees (MSTs) of all conditions. (a) MSTs represented in a circular pattern. Colors denote brain region, with black, yellow, brown, light blue, blue, and pink for frontal, subcortical, parietal, temporal, basal ganglia, and occipital regions, respectively. (b) MSTs illustrated in the corresponding brain structure. Node colors are as follows: red, blue, yellow, magenta, green, and light blue for frontal, subcortical, parietal, temporal, basal ganglia, and occipital regions, respectively. (c) Common MSTs. Common edges consisted of edges between homologous regions in bilateral hemispheres and between neighboring regions, with a few exceptions.



Supplementary Figure S4. The minimum spanning trees (MSTs) of all conditions. (a) Each MST subtracted from the common MST. (b) Common MSTs represented in a circular pattern.

국문초록

시청각 언어 처리시의 기능적 뇌 연결성: 그래프 필트레이션 방법 적용

김희정

서울대학교 대학원

협동과정 인지과학

우리가 청각 언어를 지각하고 이해함에 있어서 시각 언어 정보는 크게 기여하는 것처럼 보이지 않는다. 그러나 시끄러운 환경에서의 시각 정보의 기여는 다르다고 할 수 있다. 시청각 언어 정보를 통합하는 문제는 언어의 지각적 통합에 대한 연구 주제로 많이 다루어져 왔다. 특히 시청각 언어 정보가 일치할 때와 그렇지 않을 때의 비교는 시청각 언어 통합이 어떤 수준에서 어느 뇌 영역에서 이루어지는 지를 밝히는 데 많은 기여를 했다. 여러 연구들에서 시청각 언어 통합이 일어나는 영역을 후측 상측두엽이나 브로카 영역, 혹은 다른 영역들로 많은 영역을 거론하여 왔다. 이번 연구에서는 시청각 언어 과제를 하는 동안의 시청각 언어 처리를 담당하는 영역들 간의 기능적인 상호관계를 이해함으로써 시청각 언어가 일치할 때와 그렇지 않을 때 통합이 일어나게 되는 메커니즘을 뇌 연결성을 통

해 알아보았다. 뇌 연결성을 알아보는 방법에는 여러 가지가 있지만 최근에 그래프 이론적 접근에 기반한 네트워크 연구가 뇌의 기능적 연결성을 연구하는데 있어 수학적 및 개념적 틀을 제공함으로 인해 유용한 툴로서 이용되어 왔다. 특히 특정한 인지 과제를 하지 않는 동안의 연구가 주가 되어 왔다. 그러나 그래프 이론을 적용한 많은 연구 결과들이 적정한 수준의 역치를 정하는데 있어서 어려움이 있었다. 왜냐하면 특정한 역치를 정하게 됨으로 인해 전체적인 네트워크를 조망하기 힘들고 역치 수준을 어떻게 정하느냐에 따라서 네트워크의 결과들이 달라지기 때문이다. 이 연구에서는 특히 이러한 점을 보완하기 위해 특정한 역치를 정하지 않고 관찰 가능한 역치를 모두 조사함으로써 전체적인 네트워크를 조망하였다. 그래서 이 연구는 Persistent homology 틀에서 그래프 필트레이션 방법을 적용함으로써 역치와 상관없이 나타날 수 있는 네트워크의 특성을 시청각 언어 지각 과제를 하는 동안에 알아보았다. 시청각 언어 과제는 4가지 조건으로 구성되었다. 시청각 언어 자극을 동시에 제시하는 조건 (AV), 무의미한 시각 자극과 함께 청각 언어자극을 제시하는 조건 (A), 무의미한 청각 자극과 함께 시각 언어 자극을 제시하는 조건 (V), 그리고 무의미한 시청각 비언어 자극을 제시하는 조건 (C)으로 이루어졌다. 기능적인 뇌 네트워크를 알아보기 위해 두 가지 종류의 노드를 정의하였다. 특정 영역이 아닌 전체 영역의 기능적인 뇌 네트워크를 알아보기 위해서 템플릿을 이용하여 그들 간의

상관 행렬을 구한 후 네트워크 분석을 하였다. 그리고 언어 과제 특정적인 기능적 뇌 네트워크를 위하여 데이터의 활성화 분석결과로부터 나온 19개의 영역으로부터 네트워크 분석을 하였다. 전체 영역을 사용한 네트워크에서는 모든 조건에서 모든 영역을 하나의 연결 성분으로 나타나게 되는 시점이 AV 조건에서 가장 빠르게 나타났다. 19개의 관심 영역을 사용한 언어과제 관련 네트워크에서는 모든 19개의 영역을 하나의 성분으로 연결하게 되는 변화를 보여주는 클러스터링의 순서를 나타내는 바코드의 기울기가 조건 간에 차이를 나타내었다. 특히 AV조건에서 가장 먼저 좌측의 후측 상 측두영역과 중측두 영역, 내측 전두 영역이 가장 먼저 연결되는 것을 발견하였다. V 조건에서는 해마와 내측 전두 영역간의 연결이 먼저 이루어졌다. 또한 클러스터링의 변화를 만드는 연결을 나타내는 Minimum spanning tree (MST) 결과로부터 AV 언어 과제 동안에 좌측의 후측 상측두 영역과 전두엽의 연결이, A 조건에서는 청각 자극을 처리하는 일차 청각 영역과 전운동영역간의 연결이, V조건에서는 해마와 내측 상 전두엽의 연결이 관찰되었다. 본 연구는 언어 과제를 하는 동안의 기능적 뇌 네트워크에서 언어 특성에 따라 뇌 영역 간의 연결성이 언어 지각의 초반에 다르다는 것을 지지하고 시청각 언어 지각을 함에 있어서 여러 감각 정보를 통합하는 데에 영역간의 통합되는 시점이나 연결되는 패턴이 언어 과제마다 다르다는 것을 살펴본 연구이다. 특히 네트워크의 변화를 볼 때 입술 읽기

가 불가능한 정상인이 시각적 언어 지각을 할 때에 언어 생산과 관련이 있는 영역과 기억을 담당하는 영역의 연결성이 초반에 강하게 이루어지는 것을 볼 수 있었다. 시청각 언어의 통합이 적어도 fMRI를 통해 측정 가능한 수준에서는 각각의 특정 감각 영역으로부터 상위 수준의 연합 영역으로의 위계적인 프로세스가 일어나기보다는 비교적 시청각 자극의 통합이 초기에 일어나는 것을 알 수 있었다.

주요어: 기능적 뇌 연결성; 시청각 언어 통합; 과제관련 fMRI; 그래프 필트레이션 방법; persistent 호몰로지

학번: 2005-30932